

; Sequence 2, Application US/07942245  
; Patent No. 5639641  
; GENERAL INFORMATION:  
; APPLICANT: PEDERSEN, Jan T.  
; APPLICANT: SEARLE, Stephen M.J.  
; APPLICANT: REES, Anthony R.  
; APPLICANT: ROGUSKA, Michael A.  
; APPLICANT: GUILD, Braydon C.  
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT  
; TITLE OF INVENTION: ANTIBODIES  
; NUMBER OF SEQUENCES: 522  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; COMPUTER: HP 9000/700 Workstation  
; OPERATING SYSTEM: UNIX  
; SOFTWARE: In house  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,245  
; FILING DATE: 09-SEP-1992  
; CLASSIFICATION: 530  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 109 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-942-245-2

Query Match 62.3%; Score 33; DB 1; Length 109;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
||| |||  
Db 94 TPRFTGGG 101

RESULT 50  
US-07-942-245-10  
; Sequence 10, Application US/07942245  
; Patent No. 5639641  
; GENERAL INFORMATION:  
; APPLICANT: PEDERSEN, Jan T.  
; APPLICANT: SEARLE, Stephen M.J.  
; APPLICANT: REES, Anthony R.  
; APPLICANT: ROGUSKA, Michael A.  
; APPLICANT: GUILD, Braydon C.  
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT  
; TITLE OF INVENTION: ANTIBODIES  
; NUMBER OF SEQUENCES: 522  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; COMPUTER: HP 9000/700 Workstation  
; OPERATING SYSTEM: UNIX  
; SOFTWARE: In house

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,245  
; FILING DATE: 09-SEP-1992  
; CLASSIFICATION: 530  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 109 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-942-245-10

Query Match 62.3%; Score 33; DB 1; Length 109;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
||| |||  
Db 94 TPRFTGGG 101

Search completed: March 7, 2004, 13:08:07  
Job time : 30 secs

; CURRENT APPLICATION NUMBER: US/09/556,877  
; CURRENT FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0  
; SEQ ID NO 179  
; LENGTH: 1776  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; US-09-556-877-179

Query Match 64.2%; Score 34; DB 4; Length 1776;  
Best Local Similarity 60.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
|| : ||||:  
Db 825 TPTLIGGGAI 834

## RESULT 45

US-09-620-412C-179  
; Sequence 179, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0  
; SEQ ID NO 179  
; LENGTH: 1776  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; US-09-620-412C-179

Query Match 64.2%; Score 34; DB 4; Length 1776;  
Best Local Similarity 60.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
|| : ||||:  
Db 825 TPTLIGGGAI 834

## RESULT 46

US-09-598-419-179  
; Sequence 179, Application US/09598419  
; Patent No. 6565856  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
; FILE REFERENCE: 210121.469C6  
; CURRENT APPLICATION NUMBER: US/09/598,419  
; CURRENT FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 357  
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0  
; SEQ ID NO 179  
; LENGTH: 1776  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; US-09-598-419-179

Query Match 64.2%; Score 34; DB 4; Length 1776;  
Best Local Similarity 60.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
|| : ||||:  
Db 825 TPTLIGGGAI 834

Db 825 TPTLIGGGAI 834

## RESULT 47

US-09-240-274-38  
; Sequence 38, Application US/09240274  
; Patent No. 6255455  
; GENERAL INFORMATION:  
; APPLICANT: Siegel, Donald L.  
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL  
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF  
; FILE REFERENCE: 09596-42U2  
; CURRENT APPLICATION NUMBER: US/09/240,274  
; CURRENT FILING DATE: 1999-01-29  
; EARLIER APPLICATION NUMBER: 60/081,380  
; EARLIER FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/028,550  
; EARLIER FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: anti-Rh(D) chain I07  
; US-09-240-274-38

Query Match 62.3%; Score 33; DB 3; Length 107;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8  
|| : ||||:  
Db 93 TPRTFGGG 100

## RESULT 48

US-09-240-274-39  
; Sequence 39, Application US/09240274  
; Patent No. 6255455  
; GENERAL INFORMATION:  
; APPLICANT: Siegel, Donald L.  
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL  
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF  
; FILE REFERENCE: 09596-42U2  
; CURRENT APPLICATION NUMBER: US/09/240,274  
; CURRENT FILING DATE: 1999-01-29  
; EARLIER APPLICATION NUMBER: 60/081,380  
; EARLIER FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/028,550  
; EARLIER FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: anti-Rh(D) chain I08  
; US-09-240-274-39

Query Match 62.3%; Score 33; DB 3; Length 107;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8  
|| : ||||:  
Db 93 TPRTFGGG 100

## RESULT 49

US-07-942-245-2

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
NUMBER OF SEQUENCES: 350  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/072,596

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 199:

SEQUENCE CHARACTERISTICS:

LENGTH: 943 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-072-596-199

Query Match 64.2%; Score 34; DB 4; Length 943;  
Best Local Similarity 66.7%; Pred. No. 7.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10

Db 47 PTVAGGGAL 55

#### RESULT 42

US-09-477-135A-131

Sequence 131, Application US/09477135A

Patent No. 6572865

GENERAL INFORMATION:

APPLICANT: Nano, Francis

TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding

TITLE OF INVENTION: immunostimulatory Peptides

FILE REFERENCE: 52888

CURRENT APPLICATION NUMBER: US/09/477,135A

CURRENT FILING DATE: 2000-01-03

PRIOR APPLICATION NUMBER: 08990823

PRIOR FILING DATE: 1997-12-15

PRIOR APPLICATION NUMBER: US 96/10375

PRIOR FILING DATE: 1996-06-14

PRIOR APPLICATION NUMBER: 60/000,254

PRIOR FILING DATE: 1995-06-15

NUMBER OF SEQ ID NOS: 169

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 131

LENGTH: 943

TYPE: PRT

ORGANISM: Mycobacterium tuberculosis

US-09-477-135A-131

Query Match 64.2%; Score 34; DB 4; Length 943;  
Best Local Similarity 66.7%; Pred. No. 7.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10

Db 47 PTVAGGGAL 55

TUBERCULOSIS

Db 154 PTVAGGGAL 162

#### RESULT 43

US-09-072-967-204

Sequence 204, Application US/09072967

Patent No. 6592877

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonio

APPLICANT: Houghton, Raymond

APPLICANT: Vedvick, Thomas S.

APPLICANT: Twardzik, Daniel R.

APPLICANT: Lodes, Michael J.

APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/072,967

FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.411C9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 204:

SEQUENCE CHARACTERISTICS:

LENGTH: 943 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-09-072-967-204

Query Match 64.2%; Score 34; DB 4; Length 943;

Best Local Similarity 66.7%; Pred. No. 7.6e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10

Db 47 PTVAGGGAL 55

#### RESULT 44

US-09-556-877-179

Sequence 179, Application US/09556877

Patent No. 6432916

GENERAL INFORMATION:

APPLICANT: Probst, Peter

APPLICANT: Bhatia, Ajay

APPLICANT: Skeiky, Yasir

APPLICANT: Fling, Steve

APPLICANT: Maisonneuve, Jeff

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.469C5

; APPLICANT: Skeiky, Yasir  
; APPLICANT: Fling, Steve  
; APPLICANT: Maisonneuve, Jeff  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C5  
; CURRENT APPLICATION NUMBER: US/09/556,877  
; CURRENT FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 195  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-556-877-195

Query Match 64.2%; Score 34; DB 4; Length 821;  
Best Local Similarity 60.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TPRVTGGGAM 10  
DB 812 TPTLIGGGAI 821

RESULT 38  
US-09-620-412C-195  
; Sequence 195, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 195  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-620-412C-195

Query Match 64.2%; Score 34; DB 4; Length 821;  
Best Local Similarity 60.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TPRVTGGGAM 10  
DB 812 TPTLIGGGAI 821

RESULT 39  
US-09-598-419-195  
; Sequence 195, Application US/09598419  
; Patent No. 6565856  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C6  
; CURRENT APPLICATION NUMBER: US/09/598,419  
; CURRENT FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 357  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 195  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Chlamydia  
US-09-598-419-195

Query Match 64.2%; Score 34; DB 4; Length 821;  
Best Local Similarity 60.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TPRVTGGGAM 10  
DB 812 TPTLIGGGAI 821

RESULT 40  
US-09-056-556-204  
; Sequence 204, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Makl, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 204:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 943 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-09-056-556-204

Query Match 64.2%; Score 34; DB 4; Length 943;  
Best Local Similarity 66.7%; Pred. No. 7.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
DB 47 PTVAGGGAL 55

RESULT 41  
US-09-072-596-199  
; Sequence 199, Application US/09072596  
; Patent No. 6458366  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonia  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 856-4000  
TELEFAX: (716) 849-0349  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-736-361A-5

Query Match 64.2%; Score 34; DB 1; Length 375;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
||| |||  
Db 341 PRVTSGGS 348

RESULT 33  
US-09-252-991A-33084  
; Sequence 33084, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 33084  
; LENGTH: 510  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-33084

Query Match 64.2%; Score 34; DB 4; Length 510;  
Best Local Similarity 75.0%; Pred. No. 4.3e-02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
||| |||  
Db 167 PRTAGGGA 174

RESULT 34  
US-09-620-412C-357  
; Sequence 357, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620,412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 357  
; LENGTH: 683  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-09-620-412C-357

Query Match 64.2%; Score 34; DB 4; Length 683;  
Best Local Similarity 60.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
||| |||  
Db 258 TPTLGGGAI 267

RESULT 35  
US-09-598-419-357  
; Sequence 357, Application US/09598419  
; Patent No. 6565856  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C6  
; CURRENT APPLICATION NUMBER: US/09/598,419  
; CURRENT FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 357  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 357  
; LENGTH: 683  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-09-598-419-357

Query Match 64.2%; Score 34; DB 4; Length 683;  
Best Local Similarity 60.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
||| |||  
Db 258 TPTLGGGAI 267

RESULT 36  
US-09-252-991A-25715  
; Sequence 25715, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25715  
; LENGTH: 717  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25715

Query Match 64.2%; Score 34; DB 4; Length 717;  
Best Local Similarity 75.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
||| |||  
Db 265 PRVTGGGA 272

RESULT 37  
US-09-556-877-195  
; Sequence 195, Application US/09556877  
; Patent No. 6432916  
; GENERAL INFORMATION:  
; APPLICANT: Probst, Peter  
; APPLICANT: Bhatia, Ajay

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24671
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24671

Query Match          56.0%; Score 35; DB 4; Length 1150;
Best Local Similarity 77.8%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TPRVTGGGA 9
      || || || || ||
Db      184 TPRTGGGA 192

RESULT 29
US-09-325-932A-71
; Sequence 71, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-325-932A-71

Query Match          64.2%; Score 34; DB 4; Length 86;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      2 PRVTGGGA 9
      || || || || ||
Db      9 PKVPGGGA 16

RESULT 30
US-09-414-828A-14
; Sequence 14, Application US/09414828A
; Patent No. 6300084
; GENERAL INFORMATION:
; APPLICANT: Drubin, David G.
; APPLICANT: Hofmann, Christian J.
; TITLE OF INVENTION: Anti-Mitotic Agents and Processes
; FILE REFERENCE: Mitotic Spindle Proteins
; CURRENT APPLICATION NUMBER: US/09/414,828A
; CURRENT FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,684
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 247

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24671
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-414-828A-14

Query Match          64.2%; Score 34; DB 4; Length 247;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 PRVTGGGA 9
      || || || || ||
Db      35 PKSTGGGA 42

RESULT 31
US-09-252-991A-24996
; Sequence 24996, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24996
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24996

Query Match          64.2%; Score 34; DB 4; Length 344;
Best Local Similarity 85.7%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 PRVTGGG 8
      || || || || ||
Db      84 PRVAGGG 90

RESULT 32
US-08-736-361A-5
; Sequence 5, Application US/08736361A
; Patent No. 5705379
; GENERAL INFORMATION:
; APPLICANT: Wilson, David B.
; APPLICANT: Lao, Guifang
; TITLE OF INVENTION: Nucleotide Sequences encoding a
; TITLE OF INVENTION: Thermostable Alkaline Protease
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One M&T Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows
; SOFTWARE: Wordperfect for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/736,361A
; FILING DATE: 23 October 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION NUMBER: 35,300
; REFERENCE/DOCKET NUMBER: 18617.0031
```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29670  
; LENGTH: 369

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29670

Query Match 66.0%; Score 35; DB 4; Length 369;

Best Local Similarity 85.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8

Db 178 PRVTGGG 184

## RESULT 25

US-09-120-817-2

; Sequence 2, Application US/09120817

; Patent No. 6172184

; GENERAL INFORMATION:

; APPLICANT: Collmer, Alan

; APPLICANT: Charkowski, Amy

; APPLICANT: Alfano, James R.

; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM

; TITLE OF INVENTION: PSEUDOMONAS SYRINGAE AND ITS USE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP

; STREET: P.O. Box 1051, Clinton Square

; CITY: Rochester

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 14603

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/120,817

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/055,107

; FILING DATE: 06-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Goldman, Michael L.

; REGISTRATION NUMBER: 30,727

; REFERENCE/DOCKET NUMBER: 19603/1741

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (716) 263-1304

; TELEFAX: (716) 263-1600

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 424 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-120-817-2

Query Match

Best Local Similarity 66.0%; Score 35; DB 3; Length 424;

; Sequence 24671, Application US/09252991A

; Patent No. 6551795

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9

Db 169 TPTATGGGS 177

## RESULT 26

US-09-431-614-14

; Sequence 14, Application US/09431614

; Patent No. 6624139

; GENERAL INFORMATION:

; APPLICANT: Wei, Zhong-Min

; APPLICANT: Schading, Richard L.

; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS

; FILE REFERENCE: 21829/41 (EBC-003)

; CURRENT APPLICATION NUMBER: US/09/431,614

; CURRENT FILING DATE: 1999-11-02

; EARLIER APPLICATION NUMBER: 60/107,243

; EARLIER FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Pseudomonas syringae

US-09-431-614-14

Query Match

Best Local Similarity 66.0%; Score 35; DB 4; Length 424;

; Sequence 24671, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27068

; LENGTH: 638

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27068

Query Match

Best Local Similarity 66.0%; Score 35; DB 4; Length 638;

; Sequence 24671, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27068

; LENGTH: 638

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27068

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27680
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27680

Query Match      69.8%; Score 37; DB 4; Length 289;
Best Local Similarity 87.5%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
      |||:||||
Db      5 PRQTGGGA 12

RESULT 21
US-09-252-991A-33082
; Sequence 33082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 33082
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33082

Query Match      67.9%; Score 36; DB 4; Length 460;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
      |||:||||
Db      321 PRUSGGGA 328

RESULT 22
US-08-941-445A-15
; Sequence 15, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 33082
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33082

Query Match      67.9%; Score 36; DB 4; Length 460;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
      |||:||||
Db      321 PRUSGGGA 328

RESULT 23
US-09-731-166-12
; Sequence 12, Application US/09731166
; Patent No. 6639126
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent J. H.
; APPLICANT: Singletary, George W.
; TITLE OF INVENTION: Production of Modified Polysaccharides
; FILE REFERENCE: 35718/206348
; CURRENT APPLICATION NUMBER: US/09/731,166
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/169,993
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 799
; TYPE: PRT
; ORGANISM: Zea mays
US-09-731-166-12

Query Match      67.9%; Score 36; DB 4; Length 799;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGG 8
      |||:||||
Db      17 PRLTGGG 23

RESULT 24
US-09-252-991A-29670
; Sequence 29670, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

```
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 799 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-941-445A-15

Query Match      67.9%; Score 36; DB 3; Length 799;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGG 8
      |||:||||
Db      17 PRLTGGG 23

RESULT 23
US-09-731-166-12
; Sequence 12, Application US/09731166
; Patent No. 6639126
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent J. H.
; APPLICANT: Singletary, George W.
; TITLE OF INVENTION: Production of Modified Polysaccharides
; FILE REFERENCE: 35718/206348
; CURRENT APPLICATION NUMBER: US/09/731,166
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/169,993
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 799
; TYPE: PRT
; ORGANISM: Zea mays
US-09-731-166-12

Query Match      67.9%; Score 36; DB 4; Length 799;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGG 8
      |||:||||
Db      17 PRLTGGG 23

RESULT 24
US-09-252-991A-29670
; Sequence 29670, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 258-5200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-468-857-2

Query Match 73.6%; Score 39; DB 2; Length 447;  
Best Local Similarity 77.8%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGAM 10  
Db 332 PRVPGGAL 340

RESULT 16  
US-09-230-421-13  
; Sequence 13, Application US/09230421  
; Patent No. 6200577  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; TITLE OF INVENTION: ANTI-HERPESVIRAL AGENTS AND ASSAYS  
; FILE REFERENCE: P18189C  
; CURRENT APPLICATION NUMBER: US/09/230,421  
; CURRENT FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDES DERIVED FROM THE VP22TRUNC  
US-09-230-421-13

Query Match 71.7%; Score 38; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGG 7  
Db 13 TPRVTGG 19

RESULT 17  
US-09-230-421-3  
; Sequence 3, Application US/09230421  
; Patent No. 6200577  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; TITLE OF INVENTION: ANTI-HERPESVIRAL AGENTS AND ASSAYS  
; FILE REFERENCE: P18189C  
; CURRENT APPLICATION NUMBER: US/09/230,421  
; CURRENT FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDES DERIVED FROM THE VP22TRUNC  
US-09-230-421-3

Query Match 71.7%; Score 38; DB 3; Length 144;

Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGG 7  
Db 138 TPRVTGG 144

RESULT 18  
US-09-252-991A-22154  
; Sequence 22154, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22154  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22154

Query Match 71.7%; Score 38; DB 4; Length 455;  
Best Local Similarity 77.8%; Pred. No. 85;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10  
Db 349 PRRTGGGAL 357

RESULT 19  
US-09-134-000C-4624  
; Sequence 4624, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4624  
; LENGTH: 125  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-4624

Query Match 69.8%; Score 37; DB 4; Length 125;  
Best Local Similarity 66.7%; Pred. No. 37;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10  
Db 1 PRRTGGGAM 9

RESULT 20  
US-09-252-991A-27680  
; Sequence 27680, Application US/09252991A  
; Patent No. 6551795

```

; ADDRESS: Biotechnology Research Institute
; STREET: 1330-A Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/310,357
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/904,075
; FILING DATE:
; APPLICATION NUMBER: EP 91.201.523.7
; FILING DATE: 18-Jun-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: William M. Blackstone
; REGISTRATION NUMBER: 29,772
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-310-357-2

```

```

Query Match 73.6%; Score 39; DB 1; Length 447;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 PRVTGGGAM 10
Db 332 PRVPGGAL 340

```

```

RESULT 14
US-08-852-2
; Sequence 2, Application US/08468852
; Patent No. 5792644
; GENERAL INFORMATION:
; APPLICANT: van den BOOGAART, Paul
; APPLICANT: KOK, Jacobus Johannus
; APPLICANT: VERMEULEN, Arnoldus Nicolaas
; TITLE OF INVENTION: COCCIDIOSIS POULTRY VACCINE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5792644el Patent Department
; STREET: 1300 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,852
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/310,357
; FILING DATE: 21-SEP-1994

```

```

; APPLICATION NUMBER: US 08/102,865
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/604,075
; FILING DATE: 18-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91.201.523.7
; FILING DATE: 18-Jun-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary E. Gormley
; REGISTRATION NUMBER: 34,409
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 447 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-852-2

```

```

Query Match 73.6%; Score 39; DB 1; Length 447;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 PRVTGGGAM 10
Db 332 PRVPGGAL 340

```

```

RESULT 15
US-08-468-857-2
; Sequence 2, Application US/08468857
; Patent No. 5925347
; GENERAL INFORMATION:
; APPLICANT: van den BOOGAART, Paul
; APPLICANT: KOK, Jacobus Johannus
; APPLICANT: VERMEULEN, Arnoldus Nicolaas
; TITLE OF INVENTION: COCCIDIOSIS POULTRY VACCINE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5925347el Patent Department
; STREET: 1300 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,857
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/310,357
; FILING DATE: 21-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,865
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/904,075
; FILING DATE: 18-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91.201.523.7
; FILING DATE: 18-Jun-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary E. Gormley
; REGISTRATION NUMBER: 34,409
; REFERENCE/DOCKET NUMBER:

```

; OTHER INFORMATION: Xaa = L, F or M  
US-09-692-170C-8

Query Match 90.6%; Score 48; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
||| |||||  
DB 1 TPRVTGGGA 9

RESULT 11  
US-08-468-853-2  
; Sequence 2, Application US/08468853  
; Patent No. 5670362  
; GENERAL INFORMATION:  
; APPLICANT: van den BOOGAART, Paul  
; APPLICANT: KOK, Jacobus Johannes  
; APPLICANT: VERMEULEN, Arnoldus Nicolaas  
; TITLE OF INVENTION: COCCIDIOSIS POULTRY VACCINE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5670362el Patent Department  
; STREET: 1300 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: U.S.A.  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,853  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/310,357  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/102,865  
; FILING DATE: 06-AUG-1993  
; APPLICATION NUMBER: US 07/604,075  
; FILING DATE: 18-JUN-1992  
; APPLICATION NUMBER: EP 91.201.523.7  
; FILING DATE: 18-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mary E. Gormley  
; REGISTRATION NUMBER: 34,409  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 258-5200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-468-853-2

Query Match 73.6%; Score 39; DB 1; Length 447;  
Best Local Similarity 77.8%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
||| |||||  
DB 332 PRVPGGAL 340

RESULT 12  
US-08-468-855-2  
; Sequence 2, Application US/08468855  
; Patent No. 5780289  
; GENERAL INFORMATION:  
; APPLICANT: van den BOOGAART, Paul  
; APPLICANT: KOK, Jacobus Johannes  
; APPLICANT: VERMEULEN, Arnoldus Nicolaas  
; TITLE OF INVENTION: COCCIDIOSIS POULTRY VACCINE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5780289el Patent Department  
; STREET: 1300 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: U.S.A.  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,855  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/310,357  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/102,865  
; FILING DATE: 06-AUG-1993  
; APPLICATION NUMBER: US 07/604,075  
; FILING DATE: 18-JUN-1992  
; APPLICATION NUMBER: EP 91.201.523.7  
; FILING DATE: 18-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mary E. Gormley  
; REGISTRATION NUMBER: 34,409  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 258-5200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 447 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-468-855-2

Query Match 73.6%; Score 39; DB 1; Length 447;  
Best Local Similarity 77.8%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
||| |||||  
DB 332 PRVPGGAL 340

RESULT 13  
US-08-310-357-2  
; Sequence 2, Application US/08310357  
; Patent No. 5789233  
; GENERAL INFORMATION:  
; APPLICANT: van den BOOGAART, Paul  
; APPLICANT: KOK, Jacobus Johannes  
; APPLICANT: VERMEULEN, Arnoldus Nicolaas  
; TITLE OF INVENTION: COCCIDIOSIS POULTRY VACCINE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Organon Teknika Corporation

```
QY      1 TPRVTGGGAM 10
DB      163 TPRVTGGGAM 172

RESULT 8
US-09-075-257A-8
; Sequence 8, Application US/09075257A
; Patent No. 6074645
; GENERAL INFORMATION:
; APPLICANT: DIAMOND, DON JEFFREY
; APPLICANT: YORK, JOANNE
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BART G. NEULAND
; STREET: 555 13TH STREET, NW SUITE 701E
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,257A
; FILING DATE: 11-MAY-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/021,298
; FILING DATE: 10-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/950,064
; FILING DATE: 14-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/747,488
; FILING DATE: 12-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NEULAND, BART G
; REGISTRATION NUMBER: 31,282
; REFERENCE/DOCKET NUMBER: 1954-112CP3
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 10
; OTHER INFORMATION: /note= "Xaa = Leu, Phe, or Met"
US-09-075-257A-8

Query Match      90.6%; Score 48; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGA 9
DB      1 TPRVTGGGA 9

RESULT 10
US-09-692-170C-8
; Sequence 8, Application US/09692170C
; Patent No. 6562345
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-346
; CURRENT APPLICATION NUMBER: US/09/692,170C
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,639
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/075,257
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: US 09/021,298
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 08/950,064
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: US 08/747,488
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variant human cytomegalovirus peptide epitope
; NAME/KEY: MISC_FEATURE
; LOCATION: (10)..(10)

; Sequence 8, Application US/09534639
; Patent No. 6251399
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J
; APPLICANT: York, Joanne
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN
; TITLE OF INVENTION: CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-343
; CURRENT APPLICATION NUMBER: US/09/534,639
; CURRENT FILING DATE: 2000-03-27
; EARLIER APPLICATION NUMBER: 09/075,257
; EARLIER FILING DATE: 1998-05-11
; EARLIER APPLICATION NUMBER: 09/021,298
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: 08/950,064
; EARLIER FILING DATE: 1997-10-14
; EARLIER APPLICATION NUMBER: 08/747,488
; EARLIER FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (10)
; OTHER INFORMATION: Xaa = Leu, Phe or Met
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human
; OTHER INFORMATION: Cytomegalovirus Epitope Variant
US-09-534-639-8

Query Match      90.6%; Score 48; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGA 9
DB      1 TPRVTGGGA 9

RESULT 9
US-09-534-639-8
; Sequence 8, Application US/09534639
; Patent No. 6251399
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J
; APPLICANT: York, Joanne
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN
; TITLE OF INVENTION: CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-343
; CURRENT APPLICATION NUMBER: US/09/534,639
; CURRENT FILING DATE: 2000-03-27
; EARLIER APPLICATION NUMBER: 09/075,257
; EARLIER FILING DATE: 1998-05-11
; EARLIER APPLICATION NUMBER: 09/021,298
; EARLIER FILING DATE: 1998-02-10
; EARLIER APPLICATION NUMBER: 08/950,064
; EARLIER FILING DATE: 1997-10-14
; EARLIER APPLICATION NUMBER: 08/747,488
; EARLIER FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variant human cytomegalovirus peptide epitope
; NAME/KEY: MISC_FEATURE
; LOCATION: (10)..(10)
```



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Howson and Howson  
;; STREET: Spring House Corporate Center, P.O. Box 457  
;; CITY: Spring House  
;; STATE: Pennsylvania  
;; COUNTRY: USA  
;; ZIP: 19477  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/171,699  
;; FILING DATE: 19-Jan-1999  
;; CLASSIFICATION: <Unknown>  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/015,717  
;; FILING DATE: 23-APR-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kodroff, Cathy A.  
;; REGISTRATION NUMBER: 33,980  
;; REFERENCE/DOCKET NUMBER: WST66APCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 215-540-9200  
;; TELEFAX: 215-540-5818  
;;  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 579 amino acids  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-171-699-6  
  
Query Match 100.0%; Score 53; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TPRVTGGGAM 10  
Db 407 TPRVTGGGAM 416  
|||||  
RESULT 6  
US-09-171-699-8  
; Sequence 8, Application US/09171699  
; Patent No. 6448389  
; GENERAL INFORMATION:  
; APPLICANT: The Wistar Institute of, Anatomy & Biology  
; Gonczol, Eva  
; Berencsi, Klara  
; Kari, Csaba  
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and  
; Uses Therefor  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Center, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171,699  
; FILING DATE: 19-Jan-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 60/015,717  
;; FILING DATE: 23-APR-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kodroff, Cathy A.  
;; REGISTRATION NUMBER: 33,980  
;; REFERENCE/DOCKET NUMBER: WST66APCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 215-540-9200  
;; TELEFAX: 215-540-5818  
;;  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 579 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-171-699-8  
  
Query Match 100.0%; Score 53; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TPRVTGGGAM 10  
Db 407 TPRVTGGGAM 416  
|||||  
RESULT 7  
US-08-605-541B-12  
; Sequence 12, Application US/08605541B  
; Patent No. 5800981  
; GENERAL INFORMATION:  
; APPLICANT: Brugseman, Catharina A.  
; APPLICANT: Vink, Cornelis  
; APPLICANT: Ramon, Albert  
; APPLICANT: Stals, Frans  
; TITLE OF INVENTION: A HUMAN CYTOMEGALOVIRUS COMBINED  
; WITH AN ANTIGEN AND ITS USE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road, Suite 100  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,541B  
; FILING DATE: February 22, 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Valeta Gregg  
; REGISTRATION NUMBER: 35,127  
; REFERENCE/DOCKET NUMBER: 07532/003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 322-5070  
; TELEFAX: (415) 854-0875  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 631 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-605-541B-12  
  
Query Match 100.0%; Score 53; DB 1; Length 631;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWMAN, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-09-075-257A-7

Query Match 100.0%; Score 53; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
| | | | | | | |  
Db 1 TPRVTGGGAM 10

## RESULT 2

US-09-534-639-7  
; Sequence 7, Application US/09534639  
; Patent No. 6251399  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J  
; APPLICANT: York, Joanne  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN  
; FILE REFERENCE: 1954-343

; CURRENT APPLICATION NUMBER: US/09/534,639  
; CURRENT FILING DATE: 2000-03-27  
; EARLIER APPLICATION NUMBER: 09/075,257  
; EARLIER FILING DATE: 1998-05-11  
; EARLIER APPLICATION NUMBER: 09/021,298  
; EARLIER FILING DATE: 1998-02-10  
; EARLIER APPLICATION NUMBER: 08/950,064  
; EARLIER FILING DATE: 1997-10-14  
; EARLIER APPLICATION NUMBER: 08/747,488  
; EARLIER FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7  
; LENGTH: 10  
; TYPE: PRT

; ORGANISM: Human Cytomegalovirus  
US-09-534-639-7

Query Match 100.0%; Score 53; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
| | | | | | | |  
Db 1 TPRVTGGGAM 10

## RESULT 3

US-09-692-170C-7  
; Sequence 7, Application US/09692170C  
; Patent No. 6562345  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS  
; FILE REFERENCE: 1954-346

; CURRENT APPLICATION NUMBER: US/09/692,170C  
; CURRENT FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 09/534,639  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/075,257  
; PRIOR FILING DATE: 1998-05-11  
; PRIOR APPLICATION NUMBER: US 09/021,298  
; PRIOR FILING DATE: 1998-02-10  
; PRIOR APPLICATION NUMBER: US 08/950,064  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: US 08/747,488  
; PRIOR FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human cytomegalovirus  
US-09-692-170C-7

Query Match 100.0%; Score 53; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
| | | | | | | |  
Db 1 TPRVTGGGAM 10

## RESULT 4

US-09-812-079A-2  
; Sequence 2, Application US/09812079A  
; Patent No. 6579970  
; GENERAL INFORMATION:  
; APPLICANT: Nicolette, Charles A.  
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS  
; FILE REFERENCE: GZ 209500

; CURRENT APPLICATION NUMBER: US/09/812,079A  
; CURRENT FILING DATE: 2001-03-19  
; PRIOR APPLICATION NUMBER: 60/191,050  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/254,989  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2  
; LENGTH: 561  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-812-079A-2

Query Match 100.0%; Score 53; DB 4; Length 561;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
| | | | | | | |  
Db 417 TPRVTGGGAM 426

## RESULT 5

US-09-171-699-6  
; Sequence 6, Application US/09171699  
; Patent No. 6448389  
; GENERAL INFORMATION:  
; APPLICANT: The Wistar Institute of, Anatomy & Biology  
; Gonczol, Eva  
; Berencsi, Klara  
; Kari, Csaba

; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and  
; Uses Therefor  
; NUMBER OF SEQUENCES: 10

333 29 54.7 749 2 US-08-568-459A-6 Sequence 6, Appli  
334 29 54.7 749 2 US-08-487-826B-6 Sequence 6, Appli  
335 29 54.7 749 4 US-09-210-288-6 Sequence 6, Appli  
336 29 54.7 754 4 US-09-252-991A-22309 Sequence 23309, A  
337 29 54.7 757 4 US-09-252-991A-23569 Sequence 23569, A  
338 29 54.7 771 4 US-09-252-991A-17243 Sequence 17243, A  
339 29 54.7 780 4 US-09-252-991A-18846 Sequence 18846, A  
340 29 54.7 785 1 US-07-841-997A-4 Sequence 4, Appli  
341 29 54.7 785 1 US-08-290-301-4 Sequence 4, Appli  
342 29 54.7 785 4 US-09-013-598-4 Sequence 4, Appli  
343 29 54.7 802 4 US-09-252-991A-25050 Sequence 25050, A  
344 29 54.7 808 4 US-09-252-991A-19470 Sequence 19470, A  
345 29 54.7 826 4 US-09-894-998A-47 Sequence 47, Appli  
346 29 54.7 840 4 US-09-252-991A-19561 Sequence 19561, A  
347 29 54.7 869 4 US-09-252-991A-22290 Sequence 22290, A  
348 29 54.7 875 4 US-09-252-991A-30056 Sequence 30056, A  
349 29 54.7 930 4 US-09-198-452A-470 Sequence 470, App  
350 29 54.7 934 4 US-09-252-991A-19515 Sequence 19515, A  
351 29 54.7 964 4 US-09-543-681A-5035 Sequence 5035, Ap  
352 29 54.7 977 4 US-09-206-942-53 Sequence 53, Appli  
353 29 54.7 979 4 US-09-543-681A-5466 Sequence 5466, Ap  
354 29 54.7 983 4 US-09-206-942-51 Sequence 51, Appli  
355 29 54.7 1000 3 US-09-352-159-25 Sequence 25, Appli  
356 29 54.7 1000 3 US-09-352-168-25 Sequence 25, Appli  
357 29 54.7 1008 4 US-09-252-991A-19329 Sequence 19329, A  
358 29 54.7 1094 4 US-09-268-347-32 Sequence 32, Appli  
359 29 54.7 1098 1 US-08-290-301-82 Sequence 82, Appli  
360 29 54.7 1201 4 US-09-013-598-82 Sequence 82, Appli  
361 29 54.7 1205 4 US-09-252-991A-32259 Sequence 32259, A  
362 29 54.7 1205 3 US-09-352-159-29 Sequence 29, Appli  
363 29 54.7 1309 4 US-09-975-413A-10 Sequence 10, Appli  
364 29 54.7 1346 2 US-08-635-121-2 Sequence 2, Appli  
365 29 54.7 1400 4 US-08-978-277A-37 Sequence 37, Appli  
366 29 54.7 1752 4 US-09-556-877-180 Sequence 180, App  
367 29 54.7 1752 4 US-09-620-412C-180 Sequence 180, App  
368 29 54.7 1752 4 US-09-598-419-180 Sequence 180, App  
369 29 54.7 1848 3 US-08-296-791-6 Sequence 6, Appli  
370 29 54.7 1848 4 US-09-839-996-6 Sequence 6, Appli  
371 29 54.7 1848 5 PCT-US95-10661A-6 Sequence 6, Appli  
372 29 54.7 2071 3 US-10-080-505-6 Sequence 6, Appli  
373 29 54.7 2071 3 US-09-415-522-6 Sequence 6, Appli  
374 29 54.7 2310 4 US-09-874-923-120 Sequence 120, App  
375 28.5 53.8 284 4 US-09-387-375-7 Sequence 7, Appli  
376 28.5 53.8 316 4 US-09-387-375-9 Sequence 9, Appli  
377 28.5 53.8 345 4 US-09-543-681A-6150 Sequence 6150, Ap  
378 28.5 53.8 412 4 US-09-489-039A-7714 Sequence 7714, Ap  
379 28.5 53.8 484 1 US-08-597-236-2 Sequence 2, Appli  
380 28.5 53.8 484 1 US-08-746-682A-2 Sequence 2, Appli  
381 28.5 53.8 654 4 US-09-252-991A-19805 Sequence 19805, A  
382 28.5 53.8 10 1 US-08-147-000B-6 Sequence 6, Appli  
383 28.5 53.8 10 1 US-08-147-000B-9 Sequence 9, Appli  
384 28.5 53.8 10 1 US-08-147-000B-12 Sequence 12, Appli  
385 28.5 53.8 11 3 US-09-206-059-16 Sequence 16, Appli  
386 28.5 53.8 13 3 US-09-206-059-40 Sequence 40, Appli  
387 28.5 53.8 14 1 US-07-854-603-18 Sequence 18, Appli  
388 28.5 53.8 14 1 US-08-147-000B-10 Sequence 10, Appli  
389 28.5 53.8 15 1 US-07-854-603-20 Sequence 20, Appli  
390 28.5 53.8 15 1 US-07-854-603-28 Sequence 28, Appli  
391 28.5 53.8 15 1 US-07-854-603-30 Sequence 30, Appli  
392 28.5 53.8 15 1 US-08-147-000B-18 Sequence 18, Appli  
393 28.5 53.8 15 4 US-09-255-501-80 Sequence 80, Appli  
394 28.5 53.8 15 4 US-09-255-501-81 Sequence 81, Appli  
395 28.5 53.8 16 1 US-07-854-603-38 Sequence 38, Appli  
396 28.5 53.8 16 1 US-07-854-603-40 Sequence 40, Appli  
397 28.5 53.8 16 2 US-08-207-481-6 Sequence 6, Appli  
398 28.5 53.8 16 5 PCT-US95-02689-6 Sequence 6, Appli  
399 28.5 53.8 17 4 US-08-817-832B-5 Sequence 5, Appli  
400 28.5 53.8 19 4 US-08-506-296B-1 Sequence 1, Appli  
401 28.5 53.8 20 1 US-07-854-603-34 Sequence 34, Appli  
402 28.5 53.8 21 6 5433940-17 Patent No. 5433940  
403 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
404 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
405 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
406 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
407 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
408 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
409 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
410 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
411 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
412 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
413 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
414 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
415 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
416 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
417 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
418 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
419 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
420 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
421 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
422 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
423 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
424 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
425 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
426 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
427 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
428 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
429 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
430 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
431 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
432 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
433 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
434 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
435 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
436 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
437 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
438 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
439 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
440 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
441 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
442 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
443 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
444 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
445 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
446 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
447 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
448 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
449 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
450 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
451 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
452 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
453 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
454 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
455 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
456 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
457 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
458 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
459 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
460 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
461 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
462 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
463 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
464 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli  
465 28.5 53.8 23 1 US-08-393-985-24 Sequence 24, Appli

## ALIGNMENTS

RESULT 1  
US-09-075-257A-7  
; Sequence 7, Application US/09075257A  
; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YORKE, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:02:32 ; Search time 23 Seconds  
(without alignments)  
22.446 Million cell updates/sec

Title: US-10-697-055-7  
Perfect score: 53  
Sequence: 1 IPRVTGGGAM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep:\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep:\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	10	3	US-09-075-257A-7
2	53	100.0	10	3	US-09-534-639-7
3	53	100.0	10	4	US-09-692-170C-7
4	53	100.0	561	4	US-09-812-079A-2
5	53	100.0	579	4	US-09-171-699-6
6	53	100.0	579	4	US-09-171-699-6
7	53	100.0	631	1	US-08-605-541B-12
8	48	90.6	10	3	US-09-075-257A-8
9	48	90.6	10	4	US-09-534-639-8
10	48	90.6	10	4	US-09-692-170C-8
11	39	73.6	447	1	US-08-468-853-2
12	39	73.6	447	1	US-08-468-855-2
13	39	73.6	447	1	US-08-310-357-2
14	39	73.6	447	1	US-08-468-853-2
15	39	73.6	447	2	US-08-468-857-2
16	38	71.7	19	3	US-09-230-421-13
17	38	71.7	144	3	US-09-230-421-3
18	38	71.7	455	4	US-09-252-991A-22154
19	37	63.8	125	4	US-09-134-000C-4624
20	37	63.8	289	4	US-09-252-991A-27680
21	36	67.9	460	4	US-09-252-991A-33082
22	36	67.9	799	4	US-08-941-445A-15
23	36	67.9	799	4	US-09-731-166-12
24	35	66.0	369	4	US-09-252-991A-29670
25	35	66.0	424	3	US-09-120-817-2
26	35	66.0	424	4	US-09-431-614-14
27	35	66.0	638	4	US-09-252-991A-27068

28	35	66.0	1150	4	US-09-252-991A-24671	Sequence 24671, A
29	34	64.2	86	4	US-09-325-932A-71	Sequence 71, Appl
30	34	64.2	247	4	US-09-414-828A-14	Sequence 14, Appl
31	34	64.2	344	4	US-09-252-991A-24996	Sequence 24996, A
32	34	64.2	375	1	US-08-736-361A-5	Sequence 5, Appl
33	34	64.2	510	4	US-09-252-991A-33084	Sequence 33084, A
34	34	64.2	683	4	US-09-620-412C-357	Sequence 357, App
35	34	64.2	683	4	US-09-598-419-357	Sequence 357, App
36	34	64.2	717	4	US-09-252-991A-25715	Sequence 25715, A
37	34	64.2	821	4	US-09-556-877-195	Sequence 195, App
38	34	64.2	821	4	US-09-620-412C-195	Sequence 195, App
39	34	64.2	821	4	US-09-598-419-195	Sequence 195, App
40	34	64.2	943	4	US-09-056-556-204	Sequence 204, App
41	34	64.2	943	4	US-09-072-596-199	Sequence 199, App
42	34	64.2	943	4	US-09-477-135A-131	Sequence 131, App
43	34	64.2	943	4	US-09-072-967-204	Sequence 204, App
44	34	64.2	1776	4	US-09-556-877-179	Sequence 179, App
45	34	64.2	1776	4	US-09-620-412C-179	Sequence 179, App
46	34	64.2	1776	4	US-09-598-419-179	Sequence 179, App
47	33	62.3	107	3	US-09-240-274-38	Sequence 38, Appl
48	33	62.3	107	3	US-09-240-274-39	Sequence 39, Appl
49	33	62.3	109	1	US-07-942-245-2	Sequence 2, Appl
50	33	62.3	109	1	US-07-942-245-10	Sequence 10, Appl
51	33	62.3	149	4	US-09-252-991A-21599	Sequence 21599, A
52	33	62.3	205	4	US-09-252-991A-29378	Sequence 29378, A
53	33	62.3	216	4	US-09-198-452A-717	Sequence 717, App
54	33	62.3	223	1	US-08-430-633-1	Sequence 1, Appl
55	33	62.3	223	1	US-08-936-854-1	Sequence 1, Appl
56	33	62.3	239	2	US-08-860-174A-2	Sequence 2, Appl
57	33	62.3	247	4	US-09-134-000C-6345	Sequence 6345, A
58	33	62.3	282	4	US-09-252-991A-32829	Sequence 32829, A
59	33	62.3	311	1	US-08-078-683A-2	Sequence 2, Appl
60	33	62.3	311	1	US-08-472-217-2	Sequence 2, Appl
61	33	62.3	311	2	US-08-488-199-6	Sequence 6, Appl
62	33	62.3	311	2	US-08-760-534A-2	Sequence 2, Appl
63	33	62.3	311	4	US-09-336-757-2	Sequence 2, Appl
64	33	62.3	311	4	US-08-471-970A-2	Sequence 2, Appl
65	33	62.3	337	4	US-09-252-991A-19602	Sequence 19602, A
66	33	62.3	406	4	US-09-851-588-6	Sequence 6, Appl
67	33	62.3	420	4	US-09-252-991A-30229	Sequence 30229, A
68	33	62.3	423	4	US-09-656-002-2	Sequence 2, Appl
69	33	62.3	435	3	US-09-008-271A-6	Sequence 6, Appl
70	33	62.3	437	4	US-09-851-588-8	Sequence 8, Appl
71	33	62.3	467	4	US-09-543-681A-6989	Sequence 6989, Ap
72	33	62.3	476	3	US-09-522-800-16	Sequence 16, Appl
73	33	62.3	476	3	US-08-924-183-1	Sequence 1, Appl
74	33	62.3	476	3	US-08-924-183-7	Sequence 7, Appl
75	33	62.3	476	4	US-09-488-364-1	Sequence 1, Appl
76	33	62.3	476	4	US-09-488-364-7	Sequence 7, Appl
77	33	62.3	476	4	US-09-113-785-1	Sequence 1, Appl
78	33	62.3	476	4	US-09-460-421-2	Sequence 2, Appl
79	33	62.3	559	2	US-08-406-855A-20	Sequence 20, Appl
80	33	62.3	559	2	US-09-206-899-20	Sequence 20, Appl
81	33	62.3	560	4	US-08-688-415-8	Sequence 8, Appl
82	33	62.3	560	4	US-09-543-681A-4606	Sequence 4606, Ap
83	33	62.3	705	4	US-08-134-001C-5356	Sequence 5356, Ap
84	33	62.3	1093	3	US-08-545-860D-55	Sequence 55, Appl
85	33	62.3	1093	5	PCT-US94-04496-55	Sequence 55, Appl
86	32	60.4	20	6	5196404-11	Patent No. 5196404
87	32	60.4	95	4	US-09-489-039A-12884	Sequence 12884, A
88	32	60.4	133	4	US-09-252-991A-22856	Sequence 22856, A
89	32	60.4	142	4	US-09-732-210-372	Sequence 372, App
90	32	60.4	201	3	US-09-220-528-116	Sequence 116, App
91	32	60.4	210	4	US-09-252-991A-16807	Sequence 16807, A
92	32	60.4	223	4	US-09-252-991A-27710	Sequence 27710, A
93	32	60.4	271	4	US-09-252-991A-28704	Sequence 28704, A
94	32	60.4	287	4	US-09-489-039A-9585	Sequence 9585, Ap
95	32	60.4	293	4	US-09-252-991A-23363	Sequence 23363, A
96	32	60.4	332	4	US-09-134-001C-4032	Sequence 4032, Ap
97	32	60.4	383	4	US-09-548-938A-11	Sequence 11, Appl
98	32	60.4	394	4	US-09-252-991A-19674	Sequence 19674, A
99	32	60.4	395	4	US-09-489-039A-8706	Sequence 8706, Ap
100	32	60.4	398	4	US-09-252-991A-26217	Sequence 26217, A

```
XX PN WO200277183-A2.
XX AC
XX DT 03-OCT-2002.
XX XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX XX
XX DR WPI: 2003-029926/02.
XX DR N-PSDB; ACA21781.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX
XX PS Claim 25; SEQ ID NO 45835; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 768 AA;

Query Match 67.9%; Score 36; DB 6; Length 768;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
   ||:|||||
Db 374 PRPGGGA 381

Search completed: March 7, 2004, 13:04:36
Job time : 81 secs
```

```
XX AAW19212;
XX AC
XX DT 10-SEP-1997 (first entry)
XX DE Corn starch branching enzyme IIb.
XX XX
XX KW Starch branching enzyme IIb; SBEIIb; corn; maize; antisense; amylopectin;
XX KW transgenic plant; pBR240.
XX OS Zea mays.
XX PN WO9722703-A2.
XX XX
XX PD 26-JUN-1997.
XX PF 12-DEC-1996; 96WO-US019678.
XX PR 20-DEC-1995; 95US-00091113P.
XX XX
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Hubbard NL, Klein TM, Broglie KE;
XX DR WPI: 1997-341694/31.
XX DR N-PSDB; AAT69729.
XX XX
XX PT Transgenic corn in which grain derived starch fine structure is
XX PT controlled - specifically branch chain distribution of amylopectin,
XX PT useful in preparation of thickened foodstuff.
XX XX
XX PS Example 1; Page 50-53; 92pp; English.
XX XX
XX CC Corn starch branching enzyme IIb (SBEIIb) (AAW19212) is an isoform of
XX CC starch branching enzyme that shows higher rates of branching with
XX CC amylopectin rather than amylose as substrate. A cDNA insert (AAT69729) in
XX CC plasmid clone pBR240 that expresses SBEIIb has been used as a starting
XX CC point in the assembly of DNA constructs (see also AAT69730, AAT69736-37)
XX CC designed to achieve suppression of SBEIIb expression in transgenic corn
XX CC plants and thereby to produce novel starches that have properties
XX CC beneficial in food and industrial applications
XX SQ Sequence 799 AA;

Query Match 67.9%; Score 36; DB 2; Length 799;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
   ||:|||||
Db 17 PRLTGGG 23

Search completed: March 7, 2004, 13:04:36
Job time : 81 secs
```

```
XX PN WO200277183-A2.
XX AC
XX DT 03-OCT-2002.
XX XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX XX
XX DR WPI: 2003-029926/02.
XX DR N-PSDB; ACA21781.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX
XX PS Claim 25; SEQ ID NO 45835; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 768 AA;

Query Match 67.9%; Score 36; DB 6; Length 768;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
   ||:|||||
Db 374 PRPGGGA 381

RESULT 50
AAW19212
ID AAW19212 standard; protein; 799 AA.
```



RESULT 47  
 ABB65619  
 ID ABB65619 standard; protein; 261 AA.  
 XX  
 AC ABB65619;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 23649.  
 XX  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW  
 XX Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 PN  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEXE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR N-PDB; ABL09722.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 PT  
 XX Disclosure; SEQ ID NO 23649; 21pp + Sequence Listing; English.  
 PS  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 261 AA;  
 SQ  
 Query Match 67.9%; Score 36; DB 4; Length 261;  
 Best Local Similarity 70.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 101 TPPVMGGGAV 110  
 ||| |||||  
 RESULT 48  
 ABP65988  
 ID ABP65988 standard; protein; 361 AA.  
 XX  
 AC ABP65988;  
 XX  
 DT 19-NOV-2002 (first entry)  
 XX  
 DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:732.  
 XX  
 DE Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;  
 KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;  
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;  
 KW  
 XX  
 KW rotavirus; food composition; pharmaceutical composition.  
 XX  
 OS Bifidobacterium longum.  
 XX  
 PN EP1227152-A1.  
 XX  
 PD 31-JUL-2002.  
 XX  
 XX 30-JAN-2001; 2001EP-00102050.  
 PF  
 XX 30-JAN-2001; 2001EP-00102050.  
 PR  
 XX (NBST ) SOC PROD NESTLE SA.  
 PA  
 XX WPI; 2002-668397/72.  
 DR  
 XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as  
 PT a probe or primer for detecting and/or identifying Bifidobacterium longum  
 PT in a biological sample.  
 PT  
 XX Claim 3; SEQ ID NO 732; 80pp; English.  
 PS  
 XX The present invention describes a polynucleotide (I) comprising a  
 CC sequence of a Bifidobacterium genome selected from the nucleotide  
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at  
 CC least 90% identity or which hybridises with the sequences given in  
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a  
 CC fusion protein, comprising a sequence selected from 1097 sequences given  
 CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a  
 CC heterologous polypeptide. (I) has antidiarrheic and antibacterial  
 CC activities, and can be used as an inhibitor of Salmonella. (I) (which is  
 CC a probe) is useful for the detection and/or identification of  
 CC Bifidobacterium longum in a biological sample. A carrier containing the  
 CC lactic acid bacterium Bifidobacterium longum NCC2705 (NCIM 1-2618) can be  
 CC used for preventing and/or treating diarrhoea brought about by pathogenic  
 CC bacteria and/or rotavirus. The carrier is a food composition selected  
 CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented  
 CC products, ice-creams, fermented cereal based products, milk based  
 CC powders, infant formula, pet food or a pharmaceutical composition  
 CC selected from tablets, liquid bacterial suspensions, dried oral  
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.  
 CC (I) is useful in DNA arrays or chips to carry out analysis of the  
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent  
 CC Bifidobacterium related nucleotide sequences given in the Sequence  
 CC Listing from the present invention but not mentioned further within the  
 CC specification. N.B. The sequence data for this patent is not represented  
 CC in the printed specification but is based on sequence information  
 CC supplied by the European Patent Office  
 XX  
 XX Sequence 361 AA;  
 SQ  
 Query Match 67.9%; Score 36; DB 5; Length 361;  
 Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 RVTGGGAM 10  
 DB 215 RCTGGGAM 222  
 ||| |||||  
 RESULT 49  
 ABU17911  
 ID ABU17911 standard; protein; 768 AA.  
 XX  
 AC ABU17911;  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #3438.  
 XX  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW  
 XX Bacillus anthracis.  
 OS



\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

68

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Db		Best Local Similarity	77.8%;	Pred. No. 87;					
	53 SPRVHGGA 61	Matches	7;	Conservative	1;	Mismatches	1;	Indels	0;
		Gaps	0;						
RESULT 43									
ABM55760									
ID	ABM55760 standard; protein; 103 AA.								
XX									
AC	ABM55760;								
XX									
DT	20-OCT-2003 (first entry)								
XX									
DE	Propionibacterium acnes predicted ORF-encoded polypeptide #20436.								
XX									
XX	Acne vulgaris; antisephorhoic; dermatological; antibacterial;								
KW	immunostimulant; immune response; vaccine.								
XX									
OS	Propionibacterium acnes.								
XX									
XX	WO2003033515-A1.								
PN									
XX									
PD	24-APR-2003.								
XX									
XX	11-OCT-2002; 2002WO-US032727.								
PF									
XX									
PR	15-OCT-2001; 2001US-00978825.								
XX									
XX	(CORI-) CORIXA CORP.								
PI	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;								
PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;								
PI	Barth B, Vallieve-Douglas J;								
XX									
XX	WPI; 2003-381789/36.								
DR	N-PSDB; ACF64530.								
XX									
XX	New Propionibacterium acnes polypeptides and polynucleotides encoding the								
PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,								
PT	or for stimulating an immune response specific for a P. acnes protein.								
XX									
XX	Example 1; SEQ ID NO 20436; 1481pp; English.								
PS									
XX	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)								
CC	encoding a Propionibacterium acnes protein. The invention also relates to								
CC	polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to								
CC	immunogenic fragments of P. acnes polypeptides. The invention								
CC	additionally encompasses expression vectors and host cells comprising a								
CC	polynucleotide of the invention; antibodies against polypeptides of the								
CC	invention; fusion proteins comprising a polypeptide of the invention; a								
CC	method for stimulating an immune response specific for a P. acnes								
CC	polypeptide and an isolated T cell population comprising T cells prepared								
CC	via this method; a vaccine composition (comprising P. acnes polypeptides,								
CC	polynucleotides, antibodies, fusion proteins, T cell populations, or								
CC	antigen-presenting cells that express the polypeptide); a method and kit								
CC	for detecting or determining the presence or absence of P. acnes in a								
CC	patient; and a method for inhibiting the development of P. acnes in a								
CC	patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion								
CC	proteins, T cell populations or antigen-presenting cells that express the								
CC	polypeptides are useful for diagnosing, preventing or treating acne								
CC	vulgaris, or for stimulating an immune response specific for a P. acnes								
CC	protein. The polynucleotides can also be used as probes or primers for								
CC	nucleic acid hybridisation. The vaccine composition is useful for the								
CC	stimulation of an immune response against P. acnes, or for treating acne,								
CC	and the kit is useful for performing a diagnostic assay. The present								
CC	sequence represents a polypeptide predicted to be encoded by an ORF (open								
CC	reading frame) contained within the P. acnes polynucleotides of the								
CC	invention. Note: The sequence data for this patent did not form part of								
CC	the printed specification, but was obtained in electronic format directly								
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences								
XX									
SQ	Sequence 103 AA;								
	Query Match	69.8%;	Score 37;	DB 6;	Length 103;				

Best Local Similarity 77.8%; Pred. No. 87;									
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
QY	1 TPRVTGGGA 9								
Db	53 SPRVHGGA 61								
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									
:									

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151085P.  
PR 27-AUG-1999; 99US-0151086P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155113P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 71.7%; Score 38; DB 3; Length 401;  
Best Local Similarity 87.5%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8

Db 80 TPSVTGGG 87  
|| |||||

## RESULT 42

AAU59241

ID AAU59241 standard; protein; 103 AA.

XX

AC AAU59241;

XX

DT 27-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #20137.

XX

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX

OS Propionibacterium acnes.

XX

PN W0200181581-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US012865.

XX

PR 21-APR-2000; 2000US-0199047P.

PR

PR 02-JUN-2000; 2000US-0208841P.

PR

PR 07-JUL-2000; 2000US-0216747P.

XX

(CORI-) CORIXA CORP.

XX

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59601.

XX

Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX

PS Example 1; SEQ ID NO 20436; 1069pp; English.

XX

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 103 AA;

Query Match 69.8%; Score 37; DB 4; Length 103;  
Best Local Similarity 77.8%; Pred. No. 87;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9

PR 29-OCT-1999; 99US-0162142P.  
Query Match 71.7%; Score 38; DB 3; Length 401;  
Best Local Similarity 87.5%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 TPRVTGGG 8  
Db 80 TPSVTGGG 87

RESULT 41  
AAG50642  
ID AAG50642 standard; protein; 401 AA.  
XX  
AC AAG50642;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 64199.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145122P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140911P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 06-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151085P.  
PR 27-AUG-1999; 99US-0151086P.  
PR 27-AUG-1999; 99US-0151087P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.

CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due to  
 CC environmental conditions, including abiotic stress such as cold, nutrient  
 CC deprivation, heat, drought, salt stress, or biotic stress such as  
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate  
 CC plant architecture, plant quality traits, plant reproduction and seed  
 CC development, endoreduplication in storage cells, storage tissues and/or  
 CC storage organs of plants or its parts. CCP is useful as an immunogen to  
 CC generate antibodies. CCP protein is useful to screen for naturally  
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP  
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to  
 CC modulate CCP activity. The present sequence represents a CCP protein of  
 CC the invention

XX SQ Sequence 383 AA;

Query Match 71.7%; Score 38; DB 5; Length 383;  
 Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
 || || || || ||  
 Db 62 TPSVTGGG 69

RESULT 39

ADD30432  
 ID ADD30432 standard; protein; 383 AA.

XX AC ADD30432;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related protein from clone G2373.

XX DE transcription factor; transgenic plant; growth rate; senescence;

XX KW seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX PN WO2003013227-A2.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-US025805.

XX PR 09-AUG-2001; 2001US-0310847P.

XX PR 19-NOV-2001; 2001US-0336049P.

XX PR 11-DEC-2001; 2001US-0338692P.

XX PR 14-JUN-2002; 2002US-00171468.

XX PA (MEND-) MENDEL BIOTECHNOLOGY INC.

XX PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

XX PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

XX PI Brown PE;

XX DR WPI; 2003-248221/24.

XX DR N-PSDB; ADD30431.

XX PT New plant transcription factor polynucleotides and polypeptides, useful  
 PT in producing transgenic plants with commercially valuable properties, rate  
 PT such as an alteration in a plant growth characteristic, e.g. growth rate  
 PT or apomixis.

XX PS Disclosure; SEQ ID NO 461; 454pp; English.

XX CC The invention relates to a number of isolated Arabidopsis thaliana cDNA  
 CC sequences and their encoded proteins which are especially transcription  
 CC factor related cDNA's and proteins. The isolated or recombinant plant  
 CC transcription factor polynucleotides and polypeptides are useful in  
 CC producing transgenic plants with commercially valuable properties, i.e.  
 CC modified or altered desirable traits as compared to a reference plant,

CC such as an alteration in a plant growth characteristic, e.g. growth rate,  
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and  
 CC flower senescence. Sequence information related to the polynucleotides  
 CC and polypeptides can also be used in bioinformatic search methods. The  
 CC transgenic plant is useful for growing a progeny plant from a parent  
 CC plant. This sequence represents one of the proteins of the invention.

XX SQ Sequence 383 AA;

Query Match 71.7%; Score 38; DB 7; Length 383;  
 Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
 || || || || ||  
 Db 62 TPSVTGGG 69

RESULT 40

AAG13980  
 ID AAG13980 standard; protein; 401 AA.

XX AC AAG13980;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 13672.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 30-APR-1999; 99US-0132407P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 06-MAY-1999; 99US-0132487P.

XX PR 07-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0134258P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 14-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.

PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 71.7%; Score 38; DB 3; Length 383;  
Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8  
||| |||||  
Db 62 TPSVTGGG 69

## RESULT 38

AAU72500  
ID AAU72500 standard; protein; 383 AA.

XX AC AAU72500;

XX DT 26-FEB-2002 (first entry)

XX DE Arabidopsis partial cell cycle protein CCP19.

XX KW Cell cycle protein; CCP; cell cycle regulation; herbicide;  
XX KW Plant growth regulator; plant development; abiotic stress; biotic stress;  
XX KW nutrient deprivation; pathogen attack; crop yield; immunogen.

XX OS Arabidopsis thaliana.

XX PN WC200185946-A2.

XX PD 15-NOV-2001.

XX PF 14-MAY-2001; 2001WO-IB001307.

XX PR 12-MAY-2000; 2000US-0204045P.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX DR WPI; 2002-062249/08.

XX DR N-PSDB; AAS96290.

XX PT New cell cycle protein and nucleic acid molecule encoding it useful for  
FT regulating cell cycle progression in plants and for identifying  
FT modulators which are useful as herbicides or plant growth regulators.

XX PS Claim 34; Fig 19; 316pp; English.

XX CC The invention relates to a novel cell cycle protein (CCP) and the  
CC polynucleotides encoding them. CCP is useful for identifying a compound  
CC which modulates the activity of the polypeptide and which binds to the  
CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP  
CC nucleic acid and polypeptide molecules are useful as modulating agents in  
CC regulating cell cycle progression in plants. CCP is useful to treat  
CC disorders characterised by insufficient or excessive production of CCP  
CC protein or production of CCP protein forms which have decreased or  
CC aberrant activity. Compounds that bind to or modulate the activity of CCP  
CC polypeptide are useful as herbicides or plant growth regulators. The  
CC polynucleotide is useful for modifying cell fate, plant development,  
CC plant morphology, biochemistry and/or physiology, the length of the G1,  
CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
CC stimulation or enhancement of cell division, DNA replication, seed set,  
CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
CC initiation and/or development, nodule function, dwarfism in plants,  
CC senescence, tolerance or resistance to stress. CCP, the polynucleotide

[illegible]



PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	99US-01334941P.	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	99US-01335124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135553P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147018P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	08-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151030P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142053P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	06-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	07-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144005P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144085P.	PR	12-OCT-1999;	99US-0158369P.
PR	16-JUL-1999;	99US-0144086P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159331P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144632P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144884P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145089P.			
PR	22-JUL-1999;	99US-0145192P.			

PR 21-JUL-1999;	99US-0144814P.	PR 19-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145132P.	PR 22-OCT-1999;	99US-0160815P.
PR 22-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145931P.	PR 25-OCT-1999;	99US-0161359P.
PR 27-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 28-JUL-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161933P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 13-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 07-OCT-1999;	99US-0157865P.		
PR 08-OCT-1999;	99US-0158029P.		
PR 12-OCT-1999;	99US-0158232P.		
PR 13-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		
PR 14-OCT-1999;	99US-0159295P.		
PR 14-OCT-1999;	99US-0159329P.		
PR 14-OCT-1999;	99US-0159330P.		
PR 14-OCT-1999;	99US-0159331P.		
PR 14-OCT-1999;	99US-0159637P.		
PR 14-OCT-1999;	99US-0159638P.		

Query Match 71.7%; Score 38; DB 3; Length 331;

Best Local Similarity 87.5%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8  
 Db 10 TPRVTGGG 17

RESULT 36

AAG13981  
 ID AAG13981 standard; protein; 383 AA.

XX AAG13981;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 13673.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130449P.

XX 28-APR-1999; 99US-0130510P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

```
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 71.7%; Score 38; DB 3; Length 331;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
   |||||
Db 10 TFSVTGGG 17

RESULT 35
AAG13982
ID AAG13982 standard; protein; 331 AA.
XX
AC AAG13982;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13674.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
```

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-013256P.  
PR 14-MAY-1999; 99US-01324218P.  
PR 14-MAY-1999; 99US-01324219P.  
PR 14-MAY-1999; 99US-01324221P.  
PR 14-MAY-1999; 99US-01324370P.  
PR 18-MAY-1999; 99US-01324768P.  
PR 19-MAY-1999; 99US-01324941P.  
PR 20-MAY-1999; 99US-01325124P.  
PR 21-MAY-1999; 99US-01325353P.  
PR 24-MAY-1999; 99US-01325629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUN-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144864P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 197 AA;  
 SQ

Query Match 71.7%; Score 38; DB 4; Length 197;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
 Db 3 PRITGGG 9  
 ||:|||||

RESULT 33  
 ABR41269  
 ID ABR41269 standard; protein; 270 AA.  
 XX  
 AC ABR41269;  
 XX  
 DT 02-JUN-2003 (first entry)  
 XX  
 DE Human DITHP intracellular signalling protein.  
 XX  
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;  
 KW cancer; cell proliferative disorder; autoimmune disorder;  
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
 KW neurological disorder; gastrointestinal disorder; transport disorder;  
 KW connective tissue disorder; drug screening; proteome analysis;  
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
 KW disease model; toxicological testing; transcript imaging;  
 KW intracellular signalling.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200297031-A2.  
 XX  
 PD 05-DEC-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US010056.  
 XX  
 PR 28-MAR-2001; 2001US-0279619P.  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 29-MAR-2001; 2001US-0280068P.  
 PR 16-MAY-2001; 2001US-0291280P.  
 PR 17-MAY-2001; 2001US-0291829P.  
 PR 17-MAY-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX  
 PA (INCYTE) INCYTE GENOMICS INC.  
 XX  
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;  
 PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Uraehka ME;  
 XX  
 DR WPI; 2003-129518/12.  
 DR N-PSDB; ACC46211.  
 XX  
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying  
 PT test compound which specifically binds to a polypeptide encoded by human  
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.

XX Claim 27; SEQ ID NO 804; 591pp; English.  
 PS  
 XX  
 CC The invention relates to novel human diagnostic and therapeutic  
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded  
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to  
 CC polynucleotide sequences at least 90% identical to the dithp cDNA  
 CC sequences of the invention; recombinant vectors, host cells and  
 CC transgenic organisms comprising a dithp nucleic acid sequence; the  
 CC recombinant production of DITHP proteins; antibodies specific for DITHP  
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of  
 CC detecting dithp nucleotide and protein sequences; methods of screening  
 CC for compounds which specifically bind a DITHP protein; and methods of  
 CC assessing the toxicity of test compounds using a dithp hybridisation  
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
 CC diagnosis of a wide variety of conditions including cancer and other cell  
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic  
 CC disorders; neurological disorders; gastrointestinal disorders; transport  
 CC disorders; and connective tissue disorders. They may also be used to  
 CC screen for modulators of protein activity or gene expression. DITHP  
 CC proteins can additionally be used in analysis of the proteome of a tissue  
 CC or cell type and to induce antibodies. The dithp nucleic acids are  
 CC additionally useful in somatic or germline gene therapy of the disorders  
 CC mentioned above, as a source of antisense sequences, as a source of  
 CC probes and primers, in genotyping and identification of individuals, in  
 CC the generation of transgenic animal models of human disease or knock in  
 CC humanised animals, in toxicological testing, and in transcript imaging.  
 CC The present sequence represents a DITHP protein which has intracellular  
 CC signalling activity. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 270 AA;  
 Query Match 71.7%; Score 38; DB 6; Length 270;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGG 7  
 Db 17 TPRVTGG 23  
 |||||

RESULT 34  
 AAG50644  
 ID AAG50644 standard; protein; 331 AA.  
 XX  
 AC AAG50644;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 64201.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.



Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
Db 332 PRVGGGAL 340  
||| |||:  
||| |||:

RESULT 28  
AAB01253  
ID AAB01253 standard; protein; 10 AA.  
XX AAB01253;  
AC AAB01253;  
DT 19-DEC-2000 (first entry)  
XX Chimaeric hepatitis C virus core protein linker peptide.  
DE Hepatitis C; targeted protein production; core protein; linker.  
KW Hepatitis C.  
XX Synthetic.  
OS WO200031127-A2.  
PN 02-JUN-2000.  
XX 23-NOV-1999; 99WO-GB003898.  
XX 26-NOV-1998; 98GB-00025953.  
XX (MEDI-) MEDICAL RES COUNCIL.  
PA Hope G, McLauchlan J;  
PI WPI; 2000-400040/34.  
XX Protein comprising a lipid globule targeting sequence consisting of a  
PT hepatitis C virus core protein, useful for targeting a protein of  
PT interest to lipid globules which are subsequently secreted into animal  
PT milk.  
XX Example; Page 42; 60pp; English.  
XX The present sequence is a peptide which was used to form a linker between  
CC chimaeric proteins consisting partially of the hepatitis C virus core  
CC protein. The core protein has been shown to be localised to lipid  
CC droplets within cells. This means that fragments of the protein, in  
CC particular those comprising amino acids 125-144 or 161-166, can be used  
CC to target proteins to lipid droplets in milk. This provides a simple and  
CC efficient means of obtaining proteins of interest, using transgenic  
CC animals. Proteins of interest include those implicated in disease and  
CC antigenic polypeptides for use as vaccines  
XX Sequence 10 AA;

Query Match 71.7%; Score 38; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGG 7  
Db 4 TPRVTGG 10  
||| |||:  
||| |||:

RESULT 29  
AAB86671  
ID AAB86671 standard; peptide; 15 AA.  
XX AAB86671;  
AC AAB86671;  
DT 20-NOV-2001 (first entry)  
XX Human cytomegalovirus strain AD169 pp65 peptide fragment SEQ ID 128.

XX Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;  
KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.  
XX Human cytomegalovirus.  
XX WO200163286-A2.  
XX 30-AUG-2001.  
XX 17-FEB-2001; 2001WO-EP001773.  
XX 22-FEB-2000; 2000DE-01009341.  
XX (KERN/) KERN F.  
XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;  
XX WPI; 2001-557718/62.  
XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and  
XX the determination of the response to the antigen comprises using a  
XX synthetic peptide library of the specific antigen.  
XX Example; Fig 2; 85pp; German.  
XX This sequence represents a novel method for stimulating CD8 or CD4 T  
CC cells for the immuno-stimulation of mammals and the determination of the  
CC response to an antigen (I). The method comprises (i) dividing the amino  
CC acid sequence of the antigen into protein fragments (II) of at least 9  
CC amino acids, whereby adjacent or neighbouring fragments are in the whole  
CC antigen sequence; (ii) synthesizing a peptide library containing (II);  
CC and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension  
CC comprising (II) in a single culture vessel. The method is used to immuno-  
CC stimulate T cells of mammals, particularly humans, and is also useful for  
CC diagnostic purposes to determine if a mammal, especially human, has  
CC previously immunologically responded to a specific protein, and to  
CC determine strength of that response. The method is suitable for in vivo  
CC or in vitro immuno-stimulation of mammalian, more preferably human T  
CC lymphocytes, where the stimulated T lymphocytes are expanded and can be  
CC transfused into a patient. AAB86544-AAB86803 represent fragments of the  
CC human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are  
CC used to illustrate the method of the invention  
XX Sequence 15 AA;

Query Match 71.7%; Score 38; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGG 7  
Db 9 TPRVTGG 15  
||| |||:  
||| |||:

RESULT 30  
AAW47205  
ID AAW47205 standard; peptide; 19 AA.  
XX AAW47205;  
AC AAW47205;  
DT 03-JUL-1998 (first entry)  
XX HSV truncated tegument protein VP22 derived peptide J.  
DE HSV; tegument protein; VP22; UL49; antiviral agent; treatment; cold sore;  
XX genital herpes; chickenpox; shingles.  
KW Herpes simplex virus unknown type.  
XX OS Synthetic.  
XX WO9804708-A1.  
PN

XX PS New transgenic mouse comprising a disruption in a CAP1 gene, useful for preparing a pharmaceutical composition for treating epilepsy.

XX PS Disclosure; Fig 2; 54pp; English.

XX CC The invention discloses a new transgenic mouse comprising a disruption in a channel activating protease 1 (CAP1) gene. CAP1 is a serine protease and there is no native expression of endogenous CAP1 gene. Transgenic mice with a heterologous disruption in the CAP1 gene exhibit an increased susceptibility to seizures which are similar to epilepsy. Disclosed is a method for producing the transgenic mouse with a disruption in the CAP1 gene which contains a targeting construct comprising a first or second polynucleotide sequence homologous to at least a first or second portion, respectively, of the CAP1 gene and a selectable marker. Also disclosed are methods for identifying an agent that modulates susceptibility to seizure or the phenotype, for identifying a potential therapeutic agent for treating epilepsy, for evaluating a potential therapeutic agent capable of affecting a condition associated with a mutation in a CAP1 gene (e.g. a CAP1 agonist) and for determining whether an agent modulates a CAP1 serine protease. The transgenic mouse, comprising a disruption in a CAP1 gene, is useful for preparing a pharmaceutical composition for treating epilepsy. The sequence presented is the mouse CAP1 protein

XX SQ Sequence 339 AA;

Query Match 73.6%; Score 39; DB 6; Length 339;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
|:|:|:|:  
Db 43 PRITGGGS 50

RESULT 26  
AAAR29924  
ID AAR29924 standard; protein; 447 AA.  
AC AAR29924;  
XX 25-MAR-2003 (revised)  
DT 28-APR-1993 (first entry)  
DE Eimeria antigen Eam200.  
XX Eimeria; coccidiosis; chickens; fowl; vaccines; immunise; vaccinate;  
KW immunoassay; test; detection; ss.  
XX Eimeria acervulina.  
XX EP519547-A2.  
XX 23-DEC-1992.  
XX 10-JUN-1992; 92EP-00201673.  
XX 18-JUN-1991; 91EP-00201523.  
XX (ALKU ) AKZO NV.  
XX (ALKU ) AKZO NOBEL NV.  
XX Kok JJ, Van Den Boogaart P, Vermeulen AN;  
XX WPI; 1992-425765/52.  
XX N-PSDB; AAQ31997.  
XX New protein-contg. epitope(s) of Eimeria antigens - useful in  
PT anticoccidiosis vaccines and diagnosis.  
XX Claim 3; Page 23; 52pp; English.

XX CC This sequence represents the Eimeria antigen Eam200. It is useful as a

CC vaccine for protecting chickens against coccidiosis. It is also useful as an immunoassay reagent. (Updated on 25-MAR-2003 to correct PN field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 447 AA;

Query Match 73.6%; Score 39; DB 2; Length 447;  
Best Local Similarity 77.8%; Pred. No. 1.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10  
|:|:|:|:|:  
Db 332 PRVPGGGAL 340

## RESULT 27

AAW58567  
ID AAW58567 standard; protein; 447 AA.

XX AC AAW58567;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 24-SEP-1998 (first entry)

XX DE Eimeria acervulina 200kD antigen fragment.

XX KW Eimeria acervulina; vaccine; coccidiosis; poultry; avian;  
KW recombinant virus vector.

XX OS Eimeria acervulina.

XX FN EP838522-A2.

XX PD 29-APR-1998.

XX PF 10-JUN-1992; 97EP-00203394.

XX PR 18-JUN-1991; 91EP-00201523.

XX PR 10-JUN-1992; 92EP-00201673.

XX (ALKU ) AKZO NOBEL NV.

XX PI Van Den Boogaart P, Kok JJ, Vermeulen AN;

XX WPI; 1998-232628/21.

XX N-PSDB; AAV31335.

XX New proteins or fragments having immunogenic determinants of Eimeria antigens, nucleic acid, recombinant vector, host cell and antibody - used as vaccine against coccidiosis in avians e.g. chickens.

Example 3; Page 23-25; 50pp; English.

XX The Eimeria acervulina 200 kD antigen, recombinant vector viruses or host cells are used to produce vaccines to protect avians against coccidiosis particularly chickens. The vaccine can also contain immunogens related to other pathogens found in poultry or may contain nucleic acid encoding them e.g. antigens of Marek's Disease virus, Newcastle Disease virus, Infectious Bronchitis virus, Infectious Bursal Disease virus, Chicken Anaemia Agent, Reo virus, Avian Retro virus, Fowl Adeno virus, Turkey Rhinotracheitis virus, E. Coli or other Eimeria species to produce a multivalent vaccine. The nucleic acids are useful as hybridisation probes to isolate sequences encoding functional variant proteins e.g. from tissues or a cDNA library from a specific Eimeria strain. The nucleic acid is additionally used in test kits for diagnosis of Eimeria infections and in immunoassays. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 447 AA;

Query Match 73.6%; Score 39; DB 2; Length 447;



XX PD 28-NOV-2002.  
 XX PF 07-FEB-2002; 2002WO-IB003040.  
 XX PR 07-FEB-2001; 2001FR-00001659.  
 XX PA (INSP ) INST PASTEUR.  
 XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 XX PI Buchrieser C;  
 XX DR WPI; 2003-148459/14.  
 XX XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX PS Claim 2; SEQ ID NO 2842; 1205pp; French.  
 XX CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of P. luminescens  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than P. luminescens and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by P.  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which P.  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated P. luminescens proteins  
 XX SQ Sequence 172 AA;  
 Query Match 81.1%; Score 43; DB 6; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 PRVTGGGA 9  
 Db 84 PRVTGGGA 91  
 RESULT 24  
 ABJ37972  
 ID ABJ37972 standard; peptide; 9 AA.  
 XX AC ABJ37972;  
 XX DT 22-MAY-2003 (first entry)  
 XX DE Human cytomegalovirus CTL epitope peptide SEQ ID No 30.  
 XX KW Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;  
 KW human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; gB; gH; IE-1;  
 KW IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;  
 KW transplantation.  
 XX OS Human cytomegalovirus.  
 XX OS Human cytomegalovirus.  
 XX FN WO2003000720-A1.  
 XX XX 03-JAN-2003.

XX 26-JUN-2002; 2002WO-AU000829.  
 XX PF 26-JUN-2001; 2001AU-00005931.  
 XX PR (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 XX PA Khanna R, Elkington RA, Walker SU;  
 XX PI WPI; 2003-300379/29.  
 XX DR New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful  
 XX PT for diagnosing, preventing or treating CMV infection, comprises pp28,  
 XX PT pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.  
 XX PS Claim 1; Page 105; 308pp; English.  
 XX CC The invention relates to a novel isolated peptide comprising one or more  
 CC cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen  
 CC of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,  
 CC pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.  
 CC The peptide comprises a sequence of about 9-20 contiguous amino acids of  
 CC the antigen. The peptide epitopes are useful in diagnosing, preventing or  
 CC treating cytomegalovirus infection in humans, and in monitoring immune  
 CC responses in various clinical settings (e.g. transplantation or  
 CC pregnancy). This sequence represents a human cytomegalovirus CTL epitope  
 CC peptide of the invention  
 XX SQ Sequence 9 AA;  
 Query Match 77.4%; Score 41; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 RVTGGGAM 10  
 Db 1 RVTGGGAM 8  
 RESULT 25  
 ABG72018  
 ID ABG72018 standard; protein; 339 AA.  
 XX AC ABG72018;  
 XX DT 24-JAN-2003 (first entry)  
 XX DE Mouse channel activating protease 1 (CAP1) protein.  
 XX KW Mouse; enzyme; transgenic; channel activating protease 1; CAP1;  
 KW serine protease; seizure; epilepsy; therapeutic; agonist; anticonvulsant.  
 XX OS Mus musculus.  
 XX FH Key Location/Qualifiers  
 XX FT Active-site 236..241  
 XX XX WO200283875-A2.  
 XX PD 24-OCT-2002.  
 XX PF 29-MAR-2002; 2002WO-US009786.  
 XX PR 29-MAR-2001; 2001US-0280509P.  
 XX PR 08-AUG-2001; 2001US-0311055P.  
 XX PR 28-MAR-2002; 2002US-00109616.  
 XX PA (DELT-) DELTAGEN INC.  
 XX PI Allen KD;  
 XX DR WPI; 2003-058636/05.  
 XX DR N-PSDB; ABS56430.

XX  
DN  
W0200294867-22

Qy 1 TPRVTGGGA 9  
 |||||  
 Db 1 TPRVTGGGA 9

RESULT 19  
 AAU05432  
 ID AAU05432 standard; peptide; 10 AA.  
 XX  
 AC AAU05432;  
 XX  
 DT 24-SEP-2001 (first entry)  
 XX  
 DE Human cytomegalovirus (HCMV) cytotoxic epitope variant #3.  
 XX  
 KW Human cytomegalovirus; HCMV; cytotoxic epitope; infection; vaccine;  
 KW cytotoxic T lymphocyte; CTL; CTL precursor; CTLp; Class I antigen;  
 KW major histocompatibility complex; MHC; cellular immune response;  
 KW CTL activator; CTLp activator; immunostimulant; variant.  
 XX  
 OS Human cytomegalovirus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 10 /label= Leu, Phe, Met  
 XX  
 PN US6251399-B1.  
 XX  
 PD 26-JUN-2001.  
 XX  
 PF 27-MAR-2000; 2000US-00534639.  
 XX  
 PR 12-NOV-1996; 96US-00747488.  
 PR 14-OCT-1997; 97US-00950064.  
 PR 10-FEB-1998; 98US-00021298.  
 PR 11-MAY-1998; 98US-00075257.  
 XX  
 PA (DIAM/) DIAMOND D J.  
 PA (YORK/) YORK J.  
 XX  
 PI Diamond DJ, York J;  
 XX  
 DR WPI; 2001-431950/46.  
 XX  
 PT New peptides that are immunogenic epitopes of the human cytomegalovirus  
 PT (CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL  
 PT precursors to elicit an immune response against human CMV by normal or  
 PT immunodeficient subjects.  
 XX  
 PS Disclosure; Col 7; 17pp; English.  
 XX  
 CC The invention relates to immunologically active peptides, and functional  
 CC variants thereof, capable of eliciting a cellular immune response to  
 CC human cytomegalovirus (HCMV) in humans. The peptides are capable of  
 CC directing human cytotoxic T lymphocytes (CTL) to recognise and lyse human  
 CC cells infected with HCMV. Such immunologically active peptides, in  
 CC association with a major histocompatibility complex (MHC) Class I  
 CC molecule, are recognised by CTLs of individuals having a latent  
 CC (inactive) HCMV infection. Vaccines comprising these peptides are useful  
 CC for activating CTLs and CTL precursors (CTLp), particularly for eliciting  
 CC a cellular immune response against HCMV by normal and immunodeficient  
 CC subjects. The peptides may be used in an assay for determining the degree  
 CC of immunostimulation caused by HCMV. The peptides may also be used for  
 CC distinguishing individuals who are seropositive from those who have not  
 CC been exposed to HCMV (seronegative individuals). The present peptide  
 CC sequence is a cytotoxic epitope variant from human cytomegalovirus  
 XX  
 SQ Sequence 10 AA;  
 Query Match 90.6%; Score 48; DB 4; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.091;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9  
 |||||  
 Db 1 TPRVTGGGA 9

RESULT 20  
 AAU10827  
 ID AAU10827 standard; peptide; 10 AA.  
 XX  
 AC AAU10827;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human cytomegalovirus pp65 protein based immunogenic epitope #4.  
 XX  
 KW Immunogenic epitope; pp65; cytotoxic T cell; CTL; HCMV; vaccine;  
 KW virulence; reactivation; coronary angioplasty; morbidity; mortality; AIDS;  
 KW acquired immunodeficiency syndrome; viral retinitis; blindness;  
 KW encephalitis; enteritis; mononucleosis; interstitial pneumonia;  
 KW malignancy; Kaposi's sarcoma; post-bone marrow transplant complication.  
 XX  
 OS Human cytomegalovirus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 10 /label= Leu, Phe, Met  
 XX  
 PN WO200172782-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PF 16-MAR-2001; 2001WO-US008576.  
 XX  
 PR 27-MAR-2000; 2000US-00534639.  
 PR 20-OCT-2000; 2000US-00692170.  
 XX  
 PA (CITY) CITY OF HOPE.  
 XX  
 PI Diamond DJ;  
 XX  
 DR WPI; 2002-041233/05.  
 XX  
 PT New immunogenic peptides useful as a vaccination against cytomegalovirus  
 PT infection, comprises epitopes of human cytomegalovirus which are  
 PT recognized by human cytotoxic T-lymphocytes.  
 XX  
 PS Claim 12; Page 10; 65pp; English.  
 XX  
 CC The invention relates to immunogenic peptides derived from human  
 CC cytomegalovirus (HCMV) which are recognised by human cytotoxic T-  
 CC lymphocytes (CTL). The invention also discloses a vaccine against HCMV  
 CC comprising one of the peptides, and a pharmaceutically acceptable  
 CC carrier, a cellular vaccine against HCMV comprising antigen presenting  
 CC cells that have been treated in vitro so that they express the peptide,  
 CC and a recombinant viral vector vaccine expressing a gene encoding the  
 CC peptide. The vaccine is used to modulate the immune response to human  
 CC cytomegalovirus infection. HCMV is implicated in a number of diseases and  
 CC conditions including reactivation after coronary angioplasty.  
 CC morbidity/mortality in AIDS (acquired immunodeficiency syndrome)  
 CC patients, viral retinitis leading to blindness, encephalitis, enteritis,  
 CC mononucleosis, interstitial pneumonia, malignancy (e.g. Kaposi's  
 CC sarcoma) and post-bone marrow transplant complications. The present  
 CC sequence is an HCMV pp65 based immunogenic epitope peptide of the  
 CC invention  
 XX  
 SQ Sequence 10 AA;  
 Query Match 90.6%; Score 48; DB 5; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.091;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 17  
 AAY09315  
 ID AAY09315 standard; peptide; 10 AA.  
 XX AC AAY09315;  
 XX AC AAY09315;  
 XX DT 08-JUL-1999 (first entry)  
 XX DE Immunogenic peptide cytotoxic T lymphocyte epitope SEQ ID NO:8 of HCMV.  
 XX KW Human cytomegalovirus; HCMV; immunologically active peptide; vaccine;  
 XX KW immune response; cytotoxic T lymphocyte; CTL; immunostimulation;  
 XX KW infection; immunosuppression; bone marrow transplant; solid organ; heart;  
 XX KW AIDS.  
 XX OS Human herpesvirus 5.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Misc-difference 10 /label= Leu, Met, Phe  
 XX FT  
 XX PN WO9919349-A1.  
 XX PD 22-APR-1999.  
 XX PF 11-MAY-1998; 98WO-US009652.  
 XX PR 14-OCT-1997; 97US-00950064.  
 XX PR 10-FEB-1998; 98US-00021298.  
 XX PA (CITY) CITY OF HOPE.  
 XX PI Diamond DJ, York J;  
 XX DR WPI; 1999-277590/23.  
 XX PT Immunogenic peptide cytotoxic T lymphocyte epitopes of human  
 XX PT cytomegalovirus.  
 XX PS Claim 1; Page 49; 64pp; English.  
 XX CC The present invention describes immunologically active peptides (IAPs)  
 XX CC capable of eliciting a cellular immune response to human cytomegalovirus  
 XX CC (HCMV). The IAP can be used in a (cellular) vaccine to augment the immune  
 XX CC system response to HCMV, or to provide immunity against HCMV. The IAP  
 XX CC (cellular) vaccine can also protect an individual having a latent HCMV  
 XX CC infection from reactivation. An antigen presenting cell can be used to  
 XX CC determine the presence or absence of HCMV-infected T lymphocytes. The  
 XX CC viral vector containing IAP encoding DNA can also be used to provide  
 XX CC immunity against HCMV. The IAP can be used to prepare HCMV-reactive human  
 XX CC cytotoxic T lymphocytes. The antigen presenting cells primed with the  
 XX CC IAPs can be used as diagnostic reagents to detect immunostimulation by  
 XX CC HCMV. They can also detect active HCMV infection or exposure to HCMV.  
 XX CC HCMV can cause opportunistic infections resulting in a variety of  
 XX CC complications in, e.g. immunosuppressed patients. The IAP vaccines impart  
 XX CC immunity to bone marrow transplant recipients, solid organ recipients,  
 XX CC heart patients, AIDS patients or women of child-bearing years, without  
 XX CC the need for ex vivo expansion of HCMV-specific cytotoxic T lymphocytes  
 XX CC (which requires sophisticated laboratory setup and is highly labor-  
 XX CC intensive and costly)  
 XX SQ Sequence 10 AA;  
 Query Match 90.6%; Score 48; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.091;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
 |||||  
 Db 1 TPRVTGGGA 9

RESULT 18  
 AAB12413  
 ID AAB12413 standard; peptide; 10 AA.  
 XX AC AAB12413;  
 XX AC AAB12413;  
 XX DT 20-OCT-2000 (first entry)  
 XX DE Human cytomegalovirus immunogenic peptide SEQ ID NO:8.  
 XX KW Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;  
 XX KW infection; immune response; vaccine; immunostimulant; antiviral;  
 XX KW immunosuppressive; immunity; immunisation.  
 XX OS Human herpesvirus 5.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Misc-difference 10 /label= Leu, Phe, Met  
 XX FT  
 XX PN US6074645-A.  
 XX PD 13-JUN-2000.  
 XX PF 11-MAY-1998; 98US-00075257.  
 XX PR 12-NOV-1996; 96US-00747488.  
 XX PR 14-OCT-1997; 97US-00950064.  
 XX PR 10-FEB-1998; 98US-00021298.  
 XX PA (CITY) CITY OF HOPE.  
 XX PI York J, Diamond DJ;  
 XX DR WPI; 2000-430383/37.  
 XX PT New immunogenic peptides useful for providing immunity against  
 XX PT cytomegalovirus (CMV) infections, are capable of eliciting cellular  
 XX PT immune response to human CMV.  
 XX PS Claim 1; Col 29; 18pp; English.  
 XX CC The present invention describes a cellular vaccine (I), which elicits a  
 XX CC MHC Class I cellular immune response to human cytomegalovirus (HCMV), and  
 XX CC comprises a peptide selected from (i) - (v), provided that (i) is not  
 XX CC another peptide (vi): (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr  
 XX CC Xaa3 Glu His Pro Thr Phe Ser Gln Tyr; (iii) Phe Xaa4 Phe Pro Lys Asp Val  
 XX CC Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe  
 XX CC Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Leu, Ile, Met, Thr or Val;  
 XX CC Xaa2 = Val, Ala, Cys, Ile, Leu or Thr; Xaa3 = Ser, Thr or Leu; Xaa4 = Val  
 XX CC or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Met or Phe; and (i) is not  
 XX CC Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for  
 XX CC providing immunity against future infections by HCMV, and for augmenting  
 XX CC the immune response of an individual who is latently infected with HCMV  
 XX CC and is at risk for reactivation of HCMV infection. The peptides are also  
 XX CC useful for imparting immunity to a bone marrow transplant recipient, a  
 XX CC solid organ recipient, a heart patient, an AIDS patient or a woman of  
 XX CC child-bearing years, without the need for ex vivo expansion of HCMV-  
 XX CC specific CTL. These peptides can directly stimulate cytotoxic T  
 XX CC lymphocytes (CTL) in vitro, thus can be used in an assay to determine the  
 XX CC degree of immunostimulation caused by HCMV. These may also be used to  
 XX CC distinguish individuals who are seropositive from those who have not been  
 XX CC exposed to HCMV and in the study of the Class I antigen-processing  
 XX CC pathway for HCMV proteins  
 XX SQ Sequence 10 AA;  
 Query Match 90.6%; Score 48; DB 3; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.091;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC stimulate T cells of mammals, particularly humans, and is also useful for  
 CC diagnostic purposes to determine if a mammal, especially human, has  
 CC previously immunologically responded to a specific protein, and to  
 CC determine strength of that response. The method is suitable for in vivo  
 CC or in vitro immuno-stimulation of mammalian, more preferably human T  
 CC lymphocytes, where the stimulated T lymphocytes are expanded and can be  
 CC transfused into a patient. AAB86544-AA86803 represent fragments of the  
 CC human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are  
 CC used to illustrate the method of the invention  
 XX  
 SQ

SQ Sequence 561 AA;

Query Match 100.0%; Score 53; DB 4; Length 561;  
 Best Local Similarity 100.0%; Pred. No. 0.87;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 |||||  
 Db 417 TPRVTGGGAM 426

RESULT 15

AAW23036  
 ID AAW23036 standard; protein; 631 AA.

XX AC AAW23036;

XX AC AAW23036;

XX 23-MAR-1998 (first entry)

XX Human cytomegalovirus combined antigen.

XX Human; cytomegalovirus; combined antigen; detection; antibody; vaccine;

XX immunoglobulin A; IgA.

XX Synthetic.

XX Human cytomegalovirus.

XX WO9731117-A2.

XX 28-AUG-1997.

XX 20-FEB-1997; 97WO-EP000865.

XX 22-FEB-1996; 96US-00605541.

XX (UYMA-) UNIV MAASTRICHT.

XX Bruggeman CA, Vink C, Stals F, Ramon A;

XX WPI; 1997-435169/40.

XX N-PSDB; AAT79194.

XX Combined antigen containing parts of 3 human cytomegalovirus proteins -  
 XX useful for early detection of virus specific antibodies and in  
 XX preventative vaccines.

XX Example 4; Page 21; 26pp; English.

XX The present sequence represents a human cytomegalovirus combined antigen,  
 XX which contains amino acid sequences from at least 3 human cytomegalovirus  
 XX (hCMV) proteins and has an increased ability to bind hCMV specific  
 XX antibodies (Ab). The device is used to detect and quantify Ab,  
 XX particularly of immunoglobulin A (IgA) type, in human body fluids and  
 XX tissues. The antigen is also used in vaccines for protection against  
 XX diseases caused by hCMV and to generate Ab for hCMV detection. The  
 XX combined antigen can detect hCMV infection at an early stage (allowing  
 XX early treatment), since sensitivity and accuracy are improved by the  
 XX interaction of immunodominant regions of early formed proteins present in  
 XX a single antigen (the combine antigen's Ab binding ability is 2-3 times  
 XX greater than for single antigens). The combined antigen also avoids the  
 XX lack of antigen standardisation which limits use of conventional assay  
 XX systems

SQ Sequence 631 AA;

Query Match 100.0%; Score 53; DB 2; Length 631;  
 Best Local Similarity 100.0%; Pred. No. 0.98;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 |||||  
 Db 163 TPRVTGGGAM 172

RESULT 16

AAW48986  
 ID AAW48986 standard; peptide; 10 AA.

XX AC AAW48986;

XX 29-SEP-1998 (first entry)

XX Generic immuno-reactive peptide CTL epitope 4 of human cytomegalovirus.  
 XX Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; generic; CD8+;  
 XX cytotoxic T-lymphocyte; MHC; major histocompatibility complex; CTL.

XX Synthetic.

XX Human herpesvirus 5.

XX Key Location/Qualifiers

XX Misc-difference 10

XX /label= Leu, Phe, Met

XX WO9821233-A2.

XX 22-MAY-1998.

XX 12-NOV-1997; 97WO-US020236.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX (CITY ) CITY OF HOPE.

XX Diamond DJ, York J;

XX WPI; 1998-297862/26.

XX Immunoreactive human cytomegalovirus epitopes - useful to vaccinate

XX against infection.

XX Claim 1; Page 48; 56pp; English.

XX The present generic peptide sequence is based on the sequence given in  
 XX AAW48985. Sequence variants of the present peptide are recognised by the  
 XX CD8+ class I major histocompatibility complex (MHC) restricted cytotoxic  
 XX T-lymphocytes of patients harboring latent HCMV infection. The peptides  
 XX are claimed to be capable of activating cytotoxic T-cell lymphocytes in  
 XX the absence of active viral replication, and thus are useful for  
 XX eliciting a cellular immune response against HCMV by normal and  
 XX immunodeficient subjects. These immunological peptides can also be  
 XX formulated as vaccines which are claimed to be useful for protecting  
 XX against HCMV infection, augmenting the immune system response to a HCMV  
 XX infection or protecting against reactivation of a latent HCMV infection

SQ Sequence 10 AA;

Query Match 90.6%; Score 48; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.091;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9  
 |||||

Db 1 TPRVTGGGA 9

XX WPI; 1997-535849/49.  
 DR N-PSDB; AAT91282.  
 XX  
 PT New human cytomegalovirus DNA constructs - encode a HCMV antigen to  
 PT induce an immune response, used for the treatment and prevention of HCMV  
 PT infection.  
 XX  
 PS Example 3; Fig 5; 66pp; English.  
 XX  
 CC This sequence comprises the full-length human cytomegalovirus (HCMV)  
 CC Towne strain segment phosphoprotein pp55; the strain AD169 pp65 sequence  
 CC is given in AAW26732. The invention provides DNA molecules (see AAT91280-  
 CC 83 and AAT99436) useful for in vitro and in vivo expression of antigenic  
 CC fragments of the HCMV genome. Preferred antigens include full-length and  
 CC transmembrane-deleted fragments of gB (see AAW27274) such as gB1-680,  
 CC phosphoprotein pp65 (see AAW27276 and AAW26732), phosphoprotein pp150  
 CC (see AAW27277), and the IE-exon 4 product (see AAW27275). The DNA molecule  
 CC is preferably in the form of a plasmid and includes a regulatory sequence  
 CC for expressing the antigen in mammals, where the antigen elicits an  
 CC immune response. The products can induce HCMV-specific immune responses  
 CC including neutralising antibodies and cytotoxic T lymphocytes and can be  
 CC used for the prevention and treatment of HCMV infections  
 XX  
 SQ Sequence 551 AA;  
 Query Match 100.0%; Score 53; DB 2; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 0.85;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 Db 407 TPRVTGGGAM 416  
 |||||  
 RESULT 13  
 ID AAE10840 standard; protein; 560 AA.  
 AC AAE10840;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Cytomegalovirus antigen pp65 protein.  
 DE  
 DE Cytomegalovirus; CMV; pp65 antigen; cytostatic; virucide; cell therapy;  
 KW vaccine; viral infection; immunotherapy; gene therapy; cancer.  
 XX  
 OS Cytomegalovirus.  
 XX  
 PN WO200170766-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 19-MAR-2001; 2001WO-US008916.  
 PF  
 XX 21-MAR-2000; 2000US-0191050P.  
 PR 12-DEC-2000; 2000US-0254989P.  
 XX  
 XX (GENZ ) GENZYME CORP.  
 PA  
 XX Nicolette CA;  
 PI  
 XX WPI; 2001-596936/67.  
 DR N-PSDB; AAD18185.  
 DR  
 XX Novel synthetic antigenic compounds useful as components of antiviral  
 PT vaccines and to expand immune effector cells which are specific for viral  
 PT infections characterized by expression of cytomegalovirus antigen pp65.  
 PT  
 XX Claim 7; Page 63-64; 66pp; English.  
 PS  
 XX The invention relates to synthetic antigenic compounds which are useful

CC as components of antiviral vaccines and to expand immune effector cells  
 CC specific for viral infections characterised by expression of (CMV)  
 CC cytomegalovirus antigen pp65. Antigenic compounds are useful for  
 CC immunotherapy and also for inducing an immuneresponse in a subject. The  
 CC antigenic compound is useful for modulating immune response to the  
 CC cognate native ligands and their corresponding native proteins, and is  
 CC useful as components of anticancer vaccines. It is also used for  
 CC expanding immune effector cells that are specific for viral infections  
 CC characterised by expression of CMV antigen pp65. Furthermore, the  
 CC antigenic compound also has diagnostic applications. The DNA encoding the  
 CC or gene transcripts that are expressed in antigen presenting cells (APC),  
 CC to confirm transduction of the polynucleotides into host cells and also  
 CC in gene therapy techniques for treating cancer. The antigenic compound  
 CC immune receptor cell is also useful as vaccines. The antigenic compounds  
 CC and their DNAs are useful for preparing medicaments for diagnosis and  
 CC treatment of diseases such as cancer. The present sequence is native CMV  
 CC antigen pp65 protein  
 XX  
 SQ Sequence 560 AA;  
 Query Match 100.0%; Score 53; DB 4; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 0.87;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 Db 417 TPRVTGGGAM 426  
 |||||  
 RESULT 14  
 ID AAB86545 standard; protein; 561 AA.  
 XX  
 AC AAB86545;  
 XX  
 DT 20-NOV-2001 (first entry)  
 XX  
 DE Human cytomegalovirus strain AD169 pp65 protein.  
 XX  
 KW Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;  
 KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.  
 XX  
 OS Human cytomegalovirus.  
 XX  
 PN WO200163286-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 17-FEB-2001; 2001WO-EP001773.  
 XX  
 PR 22-FEB-2000; 2000DE-01009341.  
 XX  
 XX (KERN/) KERN F.  
 PA  
 XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatazmas E;  
 PI  
 XX WPI; 2001-557718/62.  
 DR  
 XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and  
 PT the determination of the response to the antigen comprises using a  
 PT synthetic peptide library of the specific antigen.  
 XX  
 PS Example; Page 28-29; 85pp; German.  
 XX  
 CC This sequence represents a novel method for stimulating CD8 or CD4 T  
 CC cells for the immuno-stimulation of mammals and the determination of the  
 CC response to an antigen (I). The method comprises (i) dividing the amino  
 CC acid sequence of the antigen into protein fragments (II) of at least 9  
 CC amino acids, whereby adjacent or neighbouring fragments are in the whole  
 CC antigen sequence; (ii) synthesizing a peptide library containing (II);  
 CC and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension  
 CC comprising (II) in a single culture vessel. The method is used to immuno-

DE Human cytomegalovirus CTL epitope peptide SEQ ID No 88.  
 XX Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;  
 KW human cytomegalovirus; HCMV; pp28; pp65; pp71; pp150; gB; gH; IE-1;  
 KW IE-2; US2; US3; US6; US11; UL18; cytomegalovirus infection; pregnancy;  
 KW transplantation.  
 XX Human cytomegalovirus.  
 XX OS  
 XX PN W02003000720-A1.  
 XX PD  
 XX PP 03-JAN-2003.  
 XX PP 26-JUN-2002; 2002WO-AU000829.  
 XX PR 26-JUN-2001; 2001AU-00005931.  
 XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
 XX PI Khanna R, Elkington RA, Walker SJ;  
 XX WI; 2003-300379/29.  
 XX New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful  
 FT for diagnosing, preventing or treating CMV infection, comprises pp28,  
 FT pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.  
 XX Claim 1; Page 106; 308pp; English.  
 XX The invention relates to a novel isolated peptide comprising one or more  
 CC cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen  
 CC of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,  
 CC pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.  
 CC The peptide comprises a sequence of about 9-20 contiguous amino acids of  
 CC the antigen. The peptide epitopes are useful in diagnosing, preventing or  
 CC treating cytomegalovirus infection in humans, and in monitoring immune  
 CC responses in various clinical settings (e.g. transplantation or  
 CC pregnancy). This sequence represents a human cytomegalovirus CTL epitope  
 XX peptide of the invention  
 XX SQ Sequence 20 AA;  
 Query Match 100.0%; Score 53; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 7 TPRVTGGGAM 16  
 RESULT 11  
 AAW26732  
 ID AAW26732 standard; protein; 551 AA.  
 AC AAW26732;  
 XX 27-APR-1998 (first entry)  
 DE Human cytomegalovirus tegument protein pp65.  
 DE HCMV; CMV; antigen; infection; therapy; vaccine; tegument protein;  
 KW phosphoprotein pp65.  
 XX Human cytomegalovirus AD169 strain.  
 XX Key Location/Qualifiers  
 FT Misc-difference 468  
 FT /note= "encoded by AAT"  
 XX W09740165-A1.  
 XX 30-OCT-1997.

XX 22-APR-1997; 97WO-US006866.  
 XX 23-APR-1996; 96US-0015717P.  
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 XX Gonczol E, Berencsi K, Kari C;  
 XX WI; 1997-535849/49.  
 XX N-PSDB; AAT99436.  
 XX New human cytomegalovirus DNA constructs - encode a HCMV antigen to  
 FT induce an immune response, used for the treatment and prevention of HCMV  
 FT infection.  
 XX Example 3; Fig 5; 66pp; English.  
 XX This sequence comprises the full-length human cytomegalovirus (HCMV)  
 CC AD169 strain tegument phosphoprotein pp65; the strain towne pp65 sequence  
 CC is given in AAW27276. The invention provides DNA molecules (see AAT91280-  
 CC 83 and AAT99436) useful for in vitro and in vivo expression of antigenic  
 CC fragments of the HCMV genome. Preferred antigens include full-length and  
 CC transmembrane-deleted fragments of gB (see AAW27274) such as gB1-680,  
 CC phosphoprotein pp65 (see AAW27276 and AAW26732), phosphoprotein pp150  
 CC (see AAW27277) and the IE-exon 4 product (see AAW27275). The DNA molecule  
 CC is preferably in the form of a plasmid and includes a regulatory sequence  
 CC for expressing the antigen in mammals, where the antigen elicits an  
 CC immune response. The products can induce HCMV-specific immune responses  
 CC including neutralising antibodies and cytotoxic T lymphocytes and can be  
 CC used for the prevention and treatment of HCMV infections  
 XX SQ Sequence 551 AA;  
 Query Match 100.0%; Score 53; DB 2; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 0.85;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 407 TPRVTGGGAM 416  
 RESULT 12  
 AAW27276  
 ID AAW27276 standard; protein; 551 AA.  
 AC AAW27276;  
 XX 27-APR-1998 (first entry)  
 DE Human cytomegalovirus tegument protein pp65.  
 DE HCMV; CMV; antigen; infection; therapy; vaccine; tegument protein;  
 KW phosphoprotein pp65.  
 XX Human cytomegalovirus Towne strain.  
 XX Key Location/Qualifiers  
 FT Misc-difference 468  
 FT /note= "encoded by AAT"  
 XX W09740165-A1.  
 XX 30-OCT-1997.  
 XX 22-APR-1997; 97WO-US006866.  
 XX 23-APR-1996; 96US-0015717P.  
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
 XX Gonczol E, Berencsi K, Kari C;

Query Match 100.0%; Score 53; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
DB 3 TPRVTGGGAM 12

RESULT 8  
AAB86673  
ID AAB86673 standard; peptide; 15 AA.

XX AC AAB86673;  
XX 20-NOV-2001 (first entry)  
XX Human cytomegalovirus strain AD169 pp65 peptide fragment SEQ ID 130.  
XX Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;  
XX CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.  
XX Human cytomegalovirus.  
XX WO200163286-A2.  
XX 30-AUG-2001.

PF 17-FEB-2001; 2001WO-EP001773.  
PR 22-FEB-2000; 2000DE-01009341.  
XX (KERN/) KERN F.

XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;  
XX WPI; 2001-557718/62.  
XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and  
XX the determination of the response to the antigen comprises using a  
XX synthetic peptide library of the specific antigen.

XX Example; Fig 2; 85pp; German.  
XX This sequence represents a novel method for stimulating CD8 or CD4 T  
XX cells for the immuno-stimulation of mammals and the determination of the  
XX response to an antigen (I). The method comprises (i) dividing the amino  
XX acid sequence of the antigen into protein fragments (II) of at least 9  
XX amino acids, whereby adjacent or neighbouring fragments are in the whole  
XX antigen sequence; (ii) synthesizing a peptide library containing (II);  
XX and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension  
XX comprising (II) in a single culture vessel. The method is used to immuno-  
XX stimulate T cells of mammals, particularly humans, and is also useful for  
XX diagnostic purposes to determine if a mammal, especially human, has  
XX previously immunologically responded to a specific protein, and to  
XX determine strength of that response. The method is suitable for in vivo  
XX or in vitro immuno-stimulation of mammalian, more preferably human T  
XX lymphocytes, where the stimulated T lymphocytes are expanded and can be  
XX transfused into a patient. AAB86544-AAB86803 represent fragments of the  
XX human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are  
XX used to illustrate the method of the invention

XX Sequence 15 AA;

Query Match 100.0%; Score 53; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
DB 1 TPRVTGGGAM 10

RESULT 9  
AAB86672  
ID AAB86672 standard; peptide; 15 AA.

XX AC AAB86672;  
XX 20-NOV-2001 (first entry)  
XX Human cytomegalovirus strain AD169 pp65 peptide fragment SEQ ID 129.  
XX Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;  
XX CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.  
XX Human cytomegalovirus.  
XX WO200163286-A2.  
XX 30-AUG-2001.

PF 17-FEB-2001; 2001WO-EP001773.  
PR 22-FEB-2000; 2000DE-01009341.  
XX (KERN/) KERN F.

XX Kern F, Volk H, Reinke P, Faulhaber N, Surel I, Khatamzas E;  
XX WPI; 2001-557718/62.  
XX Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and  
XX the determination of the response to the antigen comprises using a  
XX synthetic peptide library of the specific antigen.

XX Example; Fig 2; 85pp; German.

XX This sequence represents a novel method for stimulating CD8 or CD4 T  
XX cells for the immuno-stimulation of mammals and the determination of the  
XX response to an antigen (I). The method comprises (i) dividing the amino  
XX acid sequence of the antigen into protein fragments (II) of at least 9  
XX amino acids, whereby adjacent or neighbouring fragments are in the whole  
XX antigen sequence; (ii) synthesizing a peptide library containing (II);  
XX and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension  
XX comprising (II) in a single culture vessel. The method is used to immuno-  
XX stimulate T cells of mammals, particularly humans, and is also useful for  
XX diagnostic purposes to determine if a mammal, especially human, has  
XX previously immunologically responded to a specific protein, and to  
XX determine strength of that response. The method is suitable for in vivo  
XX or in vitro immuno-stimulation of mammalian, more preferably human T  
XX lymphocytes, where the stimulated T lymphocytes are expanded and can be  
XX transfused into a patient. AAB86544-AAB86803 represent fragments of the  
XX human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are  
XX used to illustrate the method of the invention

XX Sequence 15 AA;

Query Match 100.0%; Score 53; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
DB 5 TPRVTGGGAM 14

RESULT 10  
ABJ38030  
ID ABJ38030 standard; peptide; 20 AA.  
XX AC ABJ38030;  
XX 22-MAY-2003 (first entry)  
XX



CC sequence is an HCMV pp65 protein immunogenic epitope upon which vaccines  
 CC and peptides of the invention are based

XX Sequence 10 AA;

Query Match 100.0%; Score 53; DB 5; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 |||||  
 Db 1 TPRVTGGGAM 10

# RESULT 6

ID ABG66768 standard; peptide; 10 AA.

XX AC ABG66768;

XX 24-SEP-2002 (first entry)

XX Cytomegalovirus lower matrix protein pp65 HLA-B0702 epitope.

XX Beta-2 microglobulin; beta-2m; cytotoxic T lymphocyte; CTL; HLA;  
 KW human leukocyte antigen; fusion protein; epitope; cytostatic; tumour;  
 KW gastrointestinal tumour; colorectal cancer; gastro-oesophageal cancer;  
 KW liver cancer; biliary tract cancer; pancreatic cancer; vaccine;  
 KW prostatic cancer; testicular cancer; lung cancer; breast cancer;  
 KW malignant melanoma; mesothelioma; brain tumour; ovarian cancer;  
 KW uterine cancer; cervical cancer; head and neck cancer; bladder cancer;  
 KW Kaposi's sarcoma; renal carcinoma; leukaemia; lymphoma;  
 KW acquired immunodeficiency syndrome; AIDS-related lymphoma.

XX Human cytomegalovirus.

XX WO200236146-A2.

XX 10-MAY-2002.

XX 01-NOV-2001; 2001WO-GB004844.

XX 02-NOV-2000; 2000GB-00026812.

XX (ISIS-) ISIS INNOVATION LTD.

XX Tafuro S, Meier U, Mcmichael AJ, Bell JI, Layton G, Hunter M;

XX WPI; 2002-508108/54.

XX New polynucleotide capable of expressing an epitope-beta2m fusion protein  
 PT useful for generating cytotoxic T lymphocyte responses against a tumor  
 PT and in restoring antigen presentation in the tumor of a host.

XX Disclosure; Page 24; 46pp; English.

XX The invention relates to a new polynucleotide capable of expressing an  
 CC epitope-beta 2m fusion protein useful for generating cytotoxic T  
 CC lymphocyte (CTL) responses against a tumour or in restoring antigen  
 CC presentation in the tumour of a host. Also included are a polynucleotide  
 CC capable of expressing an epitope-beta-2m fusion protein in combination  
 CC with a vaccination agent that stimulates a CTL response against the  
 CC epitope of the fusion protein for simultaneous, separate or sequential  
 CC use in the treatment of cancer and a method of treating a tumour by  
 CC administering a capable of expressing an epitope-beta-2m fusion protein,  
 CC and optionally a vaccination agent that stimulates a CTL response against  
 CC the epitope of the fusion protein. The polynucleotide is useful for  
 CC generating CTL responses against tumours, for restoring antigen  
 CC presentation in the tumour, and subsequently for treating cancers, such  
 CC as gastrointestinal tumour, prostatic, testicular, lung or breast cancer,  
 CC malignant melanoma, mesothelioma, brain tumour, ovarian cancer, uterine  
 CC cancer including cervical cancer, cancer of the head and neck, bladder  
 CC cancer, Kaposi's sarcoma, AIDS (acquired immunodeficiency syndrome) -

CC related Kaposi's sarcoma, sarcomas, osteosarcoma, renal carcinoma, and  
 CC haematopoietic malignant tumours such as leukaemia and lymphoma. The  
 CC epitope is an HLA (human leukocyte antigen) peptide derived from a viral  
 CC or tumour antigen. The present sequence is a viral HLA epitope used in  
 CC the fusion proteins of the invention

XX Sequence 10 AA;

Query Match 100.0%; Score 53; DB 5; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 |||||  
 Db 1 TPRVTGGGAM 10

# RESULT 7

AAAY10218

ID AAY10218 standard; peptide; 15 AA.

XX AC AAY10218;

XX 12-MAY-1999 (first entry)

XX T cell epitope/MHC ligand SEQ ID NO:148.

XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.

XX Synthetic.

OS Human herpesvirus 5.

XX WO9902183-A2.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014289.

XX 10-JUL-1997; 97CA-02209815.

XX 10-DEC-1997; 97US-00988320.

XX (CTL-) CTL IMMUNOTHERAPIES CORP.

XX Kuendig TM, Simard JJJ;

XX WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level of  
 PT antigen in the lymphatic system of a mammal so as to provide a sustained  
 PT CTL response, used to treat, e.g. AIDS.

XX Disclosure; Page 29; 199pp; English.

XX The present invention describes a method of inducing and/or sustaining an  
 CC immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 CC method comprises: (a) delivering an antigen to the mammal at a level to  
 CC induce an immunological CTL response in the mammal; and (b) maintaining  
 CC the level of the antigen in the mammal's lymphatic system to maintain the  
 CC immunologic CTL response. The method can be used for the delivery of e.g.  
 CC a differentiation antigen, a tumour-specific multilineage antigen, an  
 CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene  
 CC antigen, or a viral antigen. They can be used for the treatment of  
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,  
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 CC to the lymphatic system provides for potent CTL stimulation that takes  
 CC place in the milieu of the lymphoid organ, and it sustains stimulation  
 CC that is necessary to keep CTL active, cytotoxic and recirculating through  
 CC the body. AAY10071 to AAY10639 represent examples of peptide antigens  
 CC given in the present invention

XX Sequence 15 AA;

CC and is at risk for reactivation of HCMV infection. The peptides are also  
 CC useful for imparting immunity to a bone marrow transplant recipient, a  
 CC solid organ recipient, a heart patient, an AIDS patient or a woman of  
 CC child-bearing years, without the need for ex vivo expansion of HCMV-  
 CC specific CTL. These peptides can directly stimulate cytotoxic T  
 CC lymphocytes (CTL) in vitro, thus can be used in an assay to determine the  
 CC degree of immunostimulation caused by HCMV. These may also be used to  
 CC distinguish individuals who are seropositive from those who have not been  
 CC exposed to HCMV and in the study of the Class I antigen-processing  
 CC pathway for HCMV proteins. The present sequence represents a peptide used  
 CC in the exemplification of the present invention

XX Sequence 10 AA;  
 SQ Query Match 100.0%; Score 53; DB 3; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 Db 1 TPRVTGGGAM 10  
 |||||

RESULT 4  
 AAE05431  
 ID AAE05431 standard; peptide; 10 AA.

XX AAE05431;

XX 24-SEP-2001 (first entry)

XX Human cytomegalovirus (HCMV) minimal cytotoxic epitope #3.

XX Human cytomegalovirus; HCMV; cytotoxic epitope; infection; vaccine;  
 KW cytotoxic T lymphocyte; CTL; CTL precursor; CTLp; Class I antigen;  
 KW major histocompatibility complex; MHC; cellular immune response;  
 KW CTL activator; CTLp activator; immunostimulant.

XX Human cytomegalovirus.

XX US6251399-B1.

XX 26-JUN-2001.

XX 27-MAR-2000; 2000US-00534639.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX 11-MAY-1998; 98US-00075257.

XX (DIAM/) DIAMOND D J.

PA (YORK/) YORK J.

XX Diamond DJ, York J;

XX WPI; 2001-431950/46.

XX New peptides that are immunogenic epitopes of the human cytomegalovirus  
 CC (CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL  
 CC precursors to elicit an immune response against human CMV by normal or  
 CC immunodeficient subjects.

XX Example 14; Col 7; 17pp; English.

XX The invention relates to immunologically active peptides, and functional  
 CC variants thereof, capable of eliciting a cellular immune response to  
 CC human cytomegalovirus (HCMV) in humans. The peptides are capable of  
 CC directing human cytotoxic T lymphocytes (CTL) to recognise and lyse human  
 CC cells infected with HCMV. Such immunologically active peptides, in  
 CC association with a major histocompatibility complex (MHC) Class I  
 CC molecule, are recognised by CTLs of individuals having a latent  
 CC (inactive) HCMV infection. Vaccines comprising these peptides are useful

CC for activating CTLs and CTL precursors (CTLp), particularly for eliciting  
 CC a cellular immune response against HCMV by normal and immunodeficient  
 CC subjects. The peptides may be used in an assay for determining the degree  
 CC of immunostimulation caused by HCMV. The peptides may also be used for  
 CC distinguishing individuals who are seropositive from those who have not  
 CC been exposed to HCMV (seronegative individuals). The present peptide  
 CC sequence is a minimal cytotoxic epitope from human cytomegalovirus. This  
 CC sequence has allele specificity to HLA B7 and subtypes

XX Sequence 10 AA;

SQ Query Match 100.0%; Score 53; DB 4; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10  
 Db 1 TPRVTGGGAM 10  
 |||||

RESULT 5  
 AAU10826  
 ID AAU10826 standard; peptide; 10 AA.

XX AAU10826;

XX 14-FEB-2002 (first entry)

XX Human cytomegalovirus pp65 protein immunogenic epitope #4.

XX Immunogenic epitope; pp65; cytotoxic T cell; CTL; HCMV; vaccine;  
 KW virucide; restenosis; coronary angioplasty; morbidity; mortality; AIDS;  
 KW acquired immunodeficiency syndrome; viral retinitis; blindness;  
 KW encephalitis; enteritis; mononucleosis; interstitial pneumonia;  
 KW malignancy; Kaposi's sarcoma; post-bone marrow transplant complication.

XX Human cytomegalovirus.

XX WO200172782-A2.

XX 04-OCT-2001.

XX 16-MAR-2001; 2001WO-US008576.

XX 27-MAR-2000; 2000US-00534639.

XX 20-OCT-2000; 2000US-00692170.

XX (CITY ) CITY OF HOPE.

XX Diamond DJ;

XX WPI; 2002-041233/05.

XX New immunogenic peptides useful as a vaccination against cytomegalovirus  
 CC infection, comprises epitopes of human cytomegalovirus which are  
 CC recognized by human cytotoxic T-lymphocytes.

XX Claim 12; Page 28; 65pp; English.

XX The invention relates to immunogenic peptides derived from human  
 CC cytomegalovirus (HCMV) which are recognised by human cytotoxic T-  
 CC lymphocytes (CTL). The invention also discloses a vaccine against HCMV  
 CC comprising one of the peptides, and a pharmaceutically acceptable  
 CC carrier, a cellular vaccine against HCMV comprising antigen presenting  
 CC cells that have been treated in vitro so that they express the peptide,  
 CC and a recombinant viral vector vaccine expressing a gene encoding the  
 CC peptide. The vaccine is used to modulate the immune response to human  
 CC cytomegalovirus infection. HCMV is implicated in a number of diseases and  
 CC conditions including restenosis after coronary angioplasty,  
 CC morbidity/mortality in AIDS (acquired immunodeficiency syndrome)  
 CC patients, viral retinitis leading to blindness, encephalitis, enteritis,  
 CC mononucleosis, interstitial pneumonia, malignancy (e.g. Kaposi's  
 CC sarcoma) and post-bone marrow transplant complications. The present

PT Immunoreactive human cytomegalovirus epitopes - useful to vaccinate  
 PT against infection.

PS Claim 9; Page 49; 56pp; English.

XX The present peptide is an immunogenic epitope which is recognised by the  
 CC CD8+ class I major histocompatibility complex (MHC) restricted cytotoxic  
 CC T-lymphocytes of patients harboring latent HCMV infection. The peptide is  
 CC claimed to be capable of activating cytotoxic T-cell lymphocytes in the  
 CC absence of active viral replication, and thus is useful for eliciting a  
 CC cellular immune response against HCMV by normal and immunodeficient  
 CC subjects. The immunological peptide can also be formulated as a vaccine  
 CC which is claimed to be useful for protecting against HCMV infection,  
 CC augmenting the immune system response to a HCMV infection or protecting  
 CC against reactivation of a latent HCMV infection

XX Sequence 10 AA;

Query Match 100.0%; Score 53; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10

Db 1 TPRVTGGGAM 10

#### RESULT 2

AAY09322  
 ID AAY09322 standard; peptide; 10 AA.

XX AC AAY09322;

XX 08-JUL-1999 (first entry)

DE Immunogenic peptide cytotoxic T lymphocyte epitope SEQ ID NO:7 of HCMV.

XX Human cytomegalovirus; HCMV; immunologically active peptide; vaccine;  
 KW immune response; cytotoxic T lymphocyte; CTL; immunostimulation;  
 KW infection; immunosuppression; bone marrow transplant; solid organ; heart;  
 KW AIDS.

XX Human herpesvirus 5.

OS Synthetic.

XX WO9919349-A1.

XX 22-APR-1999.

XX 11-MAY-1998; 98WO-US009652.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY ) CITY OF HOPE.

XX Diamond DJ, York J;

XX WPI; 1999-277590/23.

XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human  
 PT cytomegalovirus.

XX Claim 9; Page 50; 64pp; English.

XX The present invention describes immunologically active peptides (IAPs)  
 CC capable of eliciting a cellular immune response to human cytomegalovirus  
 CC (hCMV). The IAP can be used in a (cellular) vaccine to augment the immune  
 CC system response to HCMV, or to provide immunity against HCMV. The IAP  
 CC (cellular) vaccine can also protect an individual having a latent HCMV  
 CC infection from reactivation. An antigen presenting cell can be used to  
 CC determine the presence or absence of HCMV-infected T lymphocytes. The  
 CC viral vector containing IAP encoding DNA can also be used to provide

CC immunity against HCMV. The IAP can be used to prepare HCMV-reactive human  
 CC cytotoxic T lymphocytes. The antigen presenting cells primed with the  
 CC IAPs can be used as diagnostic reagents to detect immunostimulation by  
 CC HCMV. They can also detect active HCMV infection or exposure to HCMV.  
 CC HCMV can cause opportunistic infections resulting in a variety of  
 CC complications in, e.g. immunosuppressed patients. The IAP vaccines impart  
 CC immunity to bone marrow transplant recipients, solid organ recipients,  
 CC heart patients, AIDS patients or women of child-bearing years, without  
 CC the need for ex vivo expansion of HCMV-specific cytotoxic T lymphocytes  
 CC (which requires sophisticated laboratory setup and is highly labor-  
 CC intensive and costly)

XX Sequence 10 AA;

Query Match 100.0%; Score 53; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10

Db 1 TPRVTGGGAM 10

#### RESULT 3

AAB12412  
 ID AAB12412 standard; peptide; 10 AA.

XX AC AAB12412;

XX 20-OCT-2000 (first entry)

DE HCMV pp65 epitope SEQ ID NO:7.

XX Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;  
 KW infection; immune response; vaccine; immunostimulant; antiviral;  
 KW immunosuppressive; immunity; immunisation.

XX Human herpesvirus 5.

XX US6074645-A.

XX 13-JUN-2000.

XX 11-MAY-1998; 98US-00075257.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY ) CITY OF HOPE.

XX York J, Diamond DJ;

XX WPI; 2000-430383/37.

XX New immunogenic peptides useful for providing immunity against  
 PT cytomegalovirus (CMV) infections, are capable of eliciting cellular  
 PT immune response to human CMV.

XX Example 14; Col 17; 18pp; English.

XX The present invention describes a cellular vaccine (I), which elicits a  
 CC MHC Class I cellular immune response to human cytomegalovirus (HCMV), and  
 CC comprises a peptide selected from (i) - (v), provided that (i) is not  
 CC another peptide (vi); (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr  
 CC Xaa3 Glu His Pro Thr Phe Ser Gln Tyr; (iii) Phe Xaa4 Phe Pro Lys Asp Val  
 CC Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe  
 CC Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Leu, Ile, Met, Thr or Val;  
 CC Xaa2 = Val, Ala, Cys, Ile, Leu or Thr; Xaa3 = Ser, Thr or Leu; Xaa4 = Val  
 CC or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Met or Phe; and (i) is not  
 CC Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for  
 CC providing immunity against future infections by HCMV, and for augmenting  
 CC the immune response of an individual who is latently infected with HCMV

391	33	62.3	114	4	AAO08773	Aao08773 Human pol	464	33	62.3	388	3	AAO27575	Aag27575 Arabidops
392	33	62.3	126	4	AAO06612	Aao06612 Human pol	465	33	62.3	388	3	AAO08797	Aag08797 Arabidops
393	33	62.3	126	5	ABP62173	Abp62173 Human imm	466	33	62.3	388	3	AAO05673	Aag05673 Arabidops
394	33	62.3	130	2	AAO04840	Aao04840 Pectate 1	467	33	62.3	397	6	ABO00558	Novel hum
395	33	62.3	136	2	AAO04840	Aao04840 Pectate 1	468	33	62.3	399	3	AAO20924	Aag20924 Arabidops
396	33	62.3	146	7	ABO09152	Abg08028 Novel hum	469	33	62.3	401	3	AAO12960	Aag12960 Arabidops
397	33	62.3	161	4	ABG08028	Abg08028 Novel hum	470	33	62.3	401	3	ABG12960	Abg12960 Arabidops
398	33	62.3	162	2	AAO56232	Aao56232 h87-115/h	471	33	62.3	405	6	ABU04916	Abu04916 Human exp
399	33	62.3	162	4	ABU04916	Abu04916 Rice prot	472	33	62.3	405	6	ABU04916	Abu04916 Human exp
400	33	62.3	183	7	AAO08149	Aao08149 Rice prot	473	33	62.3	406	3	AAO20880	Aag20880 Arabidops
401	33	62.3	185	7	AAO08149	Aao08149 Rice prot	474	33	62.3	406	3	AAO20880	Aag20880 Arabidops
402	33	62.3	188	2	AAO66802	Aao66802 N-termina	475	33	62.3	406	3	AAO20880	Aag20880 Arabidops
403	33	62.3	201	4	ABG11352	Abg11352 Novel hum	476	33	62.3	406	3	AAO20880	Aag20880 Arabidops
404	33	62.3	205	4	ABG11352	Abg11352 Novel hum	477	33	62.3	406	3	AAO20880	Aag20880 Arabidops
405	33	62.3	211	2	AAO56231	Aao56231 h65-115/h	478	33	62.3	406	3	AAO20880	Aag20880 Arabidops
406	33	62.3	211	2	AAO56231	Aao56231 h65-115/h	479	33	62.3	406	3	AAO20880	Aag20880 Arabidops
407	33	62.3	211	2	AAO56231	Aao56231 h65-115/h	480	33	62.3	406	3	AAO20880	Aag20880 Arabidops
408	33	62.3	211	2	AAO56231	Aao56231 h65-115/h	481	33	62.3	406	3	AAO20880	Aag20880 Arabidops
409	33	62.3	216	6	ABR01487	Abro1487 Human ant	482	33	62.3	406	3	AAO20880	Aag20880 Arabidops
410	33	62.3	219	2	AAO78195	Aao78195 Human mes	483	33	62.3	406	3	AAO20880	Aag20880 Arabidops
411	33	62.3	219	2	AAO78195	Aao78195 Human mes	484	33	62.3	406	3	AAO20880	Aag20880 Arabidops
412	33	62.3	219	4	AAO98984	Aao98984 Human any	485	33	62.3	406	3	AAO20880	Aag20880 Arabidops
413	33	62.3	219	5	ABO97185	Abob97185 Novel hum	486	33	62.3	406	3	AAO20880	Aag20880 Arabidops
414	33	62.3	223	2	AAO87020	Aao87020 BZLF2 pro	487	33	62.3	406	3	AAO20880	Aag20880 Arabidops
415	33	62.3	223	2	AAO47351	Aao47351 Epstein-b	488	33	62.3	406	3	AAO20880	Aag20880 Arabidops
416	33	62.3	223	2	AAO47351	Aao47351 Epstein-b	489	33	62.3	406	3	AAO20880	Aag20880 Arabidops
417	33	62.3	228	2	AAO66804	Aao66804 N-termina	490	33	62.3	406	3	AAO20880	Aag20880 Arabidops
418	33	62.3	228	2	AAO66804	Aao66804 N-termina	491	33	62.3	406	3	AAO20880	Aag20880 Arabidops
419	33	62.3	233	2	AAO64817	Aao64817 Truncated	492	33	62.3	406	3	AAO20880	Aag20880 Arabidops
420	33	62.3	236	2	AAO22565	Aao22565 Vlyls-Huc	493	33	62.3	406	3	AAO20880	Aag20880 Arabidops
421	33	62.3	239	2	AAO09813	Aao09813 Vlyls-lin	494	33	62.3	406	3	AAO20880	Aag20880 Arabidops
422	33	62.3	239	2	AAO35561	Aao35561 EcoRI-Hin	495	33	62.3	406	3	AAO20880	Aag20880 Arabidops
423	33	62.3	239	4	AAO07498	Aao07498 Synthetic	496	33	62.3	406	3	AAO20880	Aag20880 Arabidops
424	33	62.3	240	3	AAO23166	Aao23166 Human col	497	33	62.3	406	3	AAO20880	Aag20880 Arabidops
425	33	62.3	244	2	AAO56229	Aao56229 D1.3 V-mi	498	33	62.3	406	3	AAO20880	Aag20880 Arabidops
426	33	62.3	251	4	AAO96659	Aao96659 Putative	499	33	62.3	406	3	AAO20880	Aag20880 Arabidops
427	33	62.3	251	4	AAO25808	Aao25808 Human pro	500	33	62.3	406	3	AAO20880	Aag20880 Arabidops
428	33	62.3	261	3	AAO29019	Aao29019 Arabidops							
429	33	62.3	261	3	AAO95915	Aao95915 S. pneumo							
430	33	62.3	263	2	AAO95915	Aao95915 S. pneumo							
431	33	62.3	267	5	ABO75351	Abob75351 Bacillus							
432	33	62.3	272	2	AAO21260	Aao21260 Scfv sequ							
433	33	62.3	272	3	AAO08359	Aao08359 Arabidops							
434	33	62.3	275	3	AAO08359	Aao08359 Arabidops							
435	33	62.3	278	3	AAO08359	Aao08359 Arabidops							
436	33	62.3	281	4	ABO61929	Abob61929 Drosophil							
437	33	62.3	285	2	AAO64810	Aao64810 Scfv anti							
438	33	62.3	289	2	AAO66796	Aao66796 Novel mou							
439	33	62.3	291	4	ABO69304	Abob69304 Drosophil							
440	33	62.3	305	3	AAO09845	Aao09845 Arabidops							
441	33	62.3	311	2	AAO55276	Aao55276 Syndecan							
442	33	62.3	311	2	AAO66793	Aao66793 Novel mou							
443	33	62.3	311	2	AAO87001	Aao87001 Mouse syn							
444	33	62.3	311	2	AAO47156	Aao47156 Mouse syn							
445	33	62.3	311	6	ABU09198	Abu09198 Mouse syn							
446	33	62.3	314	3	AAO08357	Aao08357 Arabidops							
447	33	62.3	317	4	AAO25633	Aao25633 Human pro							
448	33	62.3	323	3	ABO01395	Abob01395 Novel hum							
449	33	62.3	333	5	ABO53363	Abob53363 Lactococc							
450	33	62.3	335	4	ABO03159	Abob03159 Novel hum							
451	33	62.3	337	4	ABO11643	Abob11643 Novel hum							
452	33	62.3	339	3	AAO09844	Aao09844 Arabidops							
453	33	62.3	339	4	ABO66995	Abob66995 Drosophil							
454	33	62.3	339	5	ABP27900	Abp27900 Streptoco							
455	33	62.3	345	2	AAO80643	Aao80643 S. pneumo							
456	33	62.3	352	3	AAO29017	Aao29017 Arabidops							
457	33	62.3	352	3	AAO20882	Aao20882 Arabidops							
458	33	62.3	374	3	AAO12962	Aao12962 Arabidops							
459	33	62.3	374	3	AAO05905	Aao05905 Arabidops							
460	33	62.3	381	3	AAO20881	Aao20881 Arabidops							
461	33	62.3	381	3	AAO20925	Aao20925 Arabidops							
462	33	62.3	388	3	AAO17400	Aao17400 Arabidops							
463	33	62.3	388	3	AAO12961	Aao12961 Arabidops							

## ALIGNMENTS

## RESULT 1

AAW48985  
ID AAW48985 standard; peptide; 10 AA.

XX AC AAW48985;

XX 29-SEP-1998 (first entry)

XX Immuno-reactive peptide CTL epitope 3 of human cytomegalovirus.

XX Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CTL; CD8+;

XX cytotoxic T-lymphocyte; MHC; major histocompatibility complex.

XX Synthetic.

XX Human herpesvirus 5.

XX WO9821233-A2.

XX 22-MAY-1998.

XX 12-NOV-1997; 97WO-US020236.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX (CITY ) CITY OF HOPE.

XX Diamond DJ, York J;

XX WPI; 1998-297862/26.

245	34	64.2	346	7	ADCS3006	Adc53006 Novel hum	318	34	64.2	346	8	ADD87981	Adc87981 Human PRO
246	34	64.2	346	7	ADCS3360	Adc53360 Novel hum	319	34	64.2	346	8	ADD86385	Adc86385 Human PRO
247	34	64.2	346	7	ADCS6551	Adc6551 Novel hum	320	34	64.2	346	8	ADD85833	Adc85833 Human PRO
248	34	64.2	346	7	ADCS1026	Adc51026 Novel hum	321	34	64.2	346	8	ADD841380	Adc841380 Human PRO
249	34	64.2	346	7	ADCS5553	Adc6553 Human PRO	322	34	64.2	346	8	ADD823409	Adc823409 Human PRO
250	34	64.2	346	7	ADCS4651	Adc54651 Novel hum	323	34	64.2	346	8	ADD823961	Adc823961 Human PRO
251	34	64.2	346	7	ADCS3612	Adc53612 Novel hum	324	34	64.2	346	8	ADD824604	Adc824604 Human PRO
252	34	64.2	346	7	ADCS9135	Adc59135 Novel hum	325	34	64.2	346	8	ADD87429	Adc87429 Human PRO
253	34	64.2	346	7	ADCS6013	Adc56013 Novel hum	326	34	64.2	346	8	ADD89295	Adc89295 Human PRO
254	34	64.2	346	7	ADCS8583	Adc58583 Novel hum	327	34	64.2	346	8	ADD88743	Adc88743 Human PRO
255	34	64.2	346	7	ADCS14497	Adc14497 Novel hum	328	34	64.2	346	8	ADD88743	Adc88743 Human PRO
256	34	64.2	346	7	ADCS03257	Adc03257 Novel hum	329	34	64.2	346	8	ADD87966	Adc87966 Human PRO
257	34	64.2	346	7	ADCS0249	Adc0249 Novel hum	330	34	64.2	346	8	ADD87966	Adc87966 Human PRO
258	34	64.2	346	7	ADCS1854	Adc81854 Human PRO	331	34	64.2	346	8	ADD87966	Adc87966 Human PRO
259	34	64.2	346	7	ADCS9668	Adc9668 Human PRO	332	34	64.2	346	8	ADD87966	Adc87966 Human PRO
260	34	64.2	346	7	ADCS48557	Adc48557 Human PRO	333	34	64.2	346	8	ADD87966	Adc87966 Human PRO
261	34	64.2	346	7	ADCS10086	Adc10086 Human PRO	334	34	64.2	346	8	ADD87966	Adc87966 Human PRO
262	34	64.2	346	7	ADCS07496	Adc07496 Novel hum	335	34	64.2	346	8	ADD87966	Adc87966 Human PRO
263	34	64.2	346	7	ADCS04661	Adc04661 Novel hum	336	34	64.2	346	8	ADD87966	Adc87966 Human PRO
264	34	64.2	346	7	ADCS82387	Adc82387 Human PRO	337	34	64.2	346	8	ADD87966	Adc87966 Human PRO
265	34	64.2	346	7	ADCS80617	Adc80617 Human PRO	338	34	64.2	346	8	ADD87966	Adc87966 Human PRO
266	34	64.2	346	7	ADCS11124	Adc11124 Human PRO	339	34	64.2	346	8	ADD87966	Adc87966 Human PRO
267	34	64.2	346	7	ADCS10419	Adc10419 Human PRO	340	34	64.2	346	8	ADD87966	Adc87966 Human PRO
268	34	64.2	346	7	ADCS48005	Adc48005 Human PRO	341	34	64.2	346	8	ADD87966	Adc87966 Human PRO
269	34	64.2	346	7	ADCS08567	Adc08567 Novel hum	342	34	64.2	346	8	ADD87966	Adc87966 Human PRO
270	34	64.2	346	7	ADCS80065	Adc80065 Novel hum	343	34	64.2	346	8	ADD87966	Adc87966 Human PRO
271	34	64.2	346	7	ADCS06816	Adc06816 Novel hum	344	34	64.2	346	8	ADD87966	Adc87966 Human PRO
272	34	64.2	346	7	ADCS11379	Adc11379 Human PRO	345	34	64.2	346	8	ADD87966	Adc87966 Human PRO
273	34	64.2	346	7	ADCS09534	Adc09534 Human PRO	346	34	64.2	346	8	ADD87966	Adc87966 Human PRO
274	34	64.2	346	7	ADCS83063	Adc83063 Human PRO	347	34	64.2	346	8	ADD87966	Adc87966 Human PRO
275	34	64.2	346	7	ADCS68075	Adc68075 Human PRO	348	34	64.2	346	8	ADD87966	Adc87966 Human PRO
276	34	64.2	346	7	ADCS41247	Adc41247 Novel hum	349	34	64.2	346	8	ADD87966	Adc87966 Human PRO
277	34	64.2	346	7	ADCS2386	Adc2386 Human PRO	350	34	64.2	346	8	ADD87966	Adc87966 Human PRO
278	34	64.2	346	7	ADCS3126	Adc3126 Human PRO	351	34	64.2	346	8	ADD87966	Adc87966 Human PRO
279	34	64.2	346	7	ADCS3678	Adc3678 Novel hum	352	34	64.2	346	8	ADD87966	Adc87966 Human PRO
280	34	64.2	346	7	ADCS5170	Adc5170 Human PRO	353	34	64.2	346	8	ADD87966	Adc87966 Human PRO
281	34	64.2	346	7	ADCS7172	Adc7172 Human PRO	354	34	64.2	346	8	ADD87966	Adc87966 Human PRO
282	34	64.2	346	7	ADCS56128	Adc56128 Human PRO	355	34	64.2	346	8	ADD87966	Adc87966 Human PRO
283	34	64.2	346	7	ADCS1834	Adc1834 Human PRO	356	34	64.2	346	8	ADD87966	Adc87966 Human PRO
284	34	64.2	346	7	ADCS02633	Adc02633 Human PRO	357	34	64.2	346	8	ADD87966	Adc87966 Human PRO
285	34	64.2	346	7	ADCS02067	Adc02067 Novel hum	358	34	64.2	346	8	ADD87966	Adc87966 Human PRO
286	34	64.2	346	7	ADCS4249	Adc4249 Novel hum	359	34	64.2	346	8	ADD87966	Adc87966 Human PRO
287	34	64.2	346	7	ADCS4566	Adc4566 Human PRO	360	34	64.2	346	8	ADD87966	Adc87966 Human PRO
288	34	64.2	346	7	ADCS2566	Adc2566 Human PRO	361	34	64.2	346	8	ADD87966	Adc87966 Human PRO
289	34	64.2	346	7	ADCS1462	Adc1462 Human PRO	362	34	64.2	346	8	ADD87966	Adc87966 Human PRO
290	34	64.2	346	7	ADCS04076	Adc04076 Human PRO	363	34	64.2	346	8	ADD87966	Adc87966 Human PRO
291	34	64.2	346	7	ADCS2373	Adc2373 Novel hum	364	34	64.2	346	8	ADD87966	Adc87966 Human PRO
292	34	64.2	346	7	ADCS2373	Adc2373 Novel hum	365	34	64.2	346	8	ADD87966	Adc87966 Human PRO
293	34	64.2	346	7	ADCS2305	Adc2305 Human PRO	366	34	64.2	346	8	ADD87966	Adc87966 Human PRO
294	34	64.2	346	7	ADCS79529	Adc79529 Human PRO	367	34	64.2	346	8	ADD87966	Adc87966 Human PRO
295	34	64.2	346	7	ADCS2065	Adc2065 Human PRO	368	34	64.2	346	8	ADD87966	Adc87966 Human PRO
296	34	64.2	346	7	ADCS118	Adc118 Human PRO	369	34	64.2	346	8	ADD87966	Adc87966 Human PRO
297	34	64.2	346	7	ADCS17882	Adc17882 Human PRO	370	34	64.2	346	8	ADD87966	Adc87966 Human PRO
298	34	64.2	346	7	ADCS3477	Adc3477 Novel hum	371	34	64.2	346	8	ADD87966	Adc87966 Human PRO
299	34	64.2	346	7	ADCS4029	Adc4029 Novel hum	372	34	64.2	346	8	ADD87966	Adc87966 Human PRO
300	34	64.2	346	7	ADCS3118	Adc3118 Human PRO	373	34	64.2	346	8	ADD87966	Adc87966 Human PRO
301	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	374	34	64.2	346	8	ADD87966	Adc87966 Human PRO
302	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	375	34	64.2	346	8	ADD87966	Adc87966 Human PRO
303	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	376	34	64.2	346	8	ADD87966	Adc87966 Human PRO
304	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	377	34	64.2	346	8	ADD87966	Adc87966 Human PRO
305	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	378	34	64.2	346	8	ADD87966	Adc87966 Human PRO
306	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	379	34	64.2	346	8	ADD87966	Adc87966 Human PRO
307	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	380	34	64.2	346	8	ADD87966	Adc87966 Human PRO
308	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	381	34	64.2	346	8	ADD87966	Adc87966 Human PRO
309	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	382	34	64.2	346	8	ADD87966	Adc87966 Human PRO
310	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	383	34	64.2	346	8	ADD87966	Adc87966 Human PRO
311	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	384	34	64.2	346	8	ADD87966	Adc87966 Human PRO
312	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	385	34	64.2	346	8	ADD87966	Adc87966 Human PRO
313	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	386	34	64.2	346	8	ADD87966	Adc87966 Human PRO
314	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	387	34	64.2	346	8	ADD87966	Adc87966 Human PRO
315	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	388	34	64.2	346	8	ADD87966	Adc87966 Human PRO
316	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	389	34	64.2	346	8	ADD87966	Adc87966 Human PRO
317	34	64.2	346	7	ADCS1938	Adc1938 Human PRO	390	34	64.2	346	8	ADD87966	Adc87966 Human PRO

99	34	64.2	263	4	AAG90941	C. glutami	172	34	64.2	346	6	ABO43364	Novel hum
100	34	64.2	269	5	ABB48345	Listeria	173	34	64.2	346	6	ADA94413	Human pro
101	34	64.2	269	6	ABU33087	Protein e	174	34	64.2	346	6	ADA74558	Human pro
102	34	64.2	288	3	AAB59022	Breast an	175	34	64.2	346	6	ADB24791	Human pro
103	34	64.2	288	3	AG05724	Arabidops	176	34	64.2	346	6	ADA82315	Human pro
104	34	64.2	346	2	AAW85739	Polypepti	177	34	64.2	346	6	ADA75278	Human pro
105	34	64.2	346	3	AAW85739	Membrane-	178	34	64.2	346	6	ADA85356	Novel hum
106	34	64.2	346	3	AAW85739	Membrane-	179	34	64.2	346	6	ADA84804	Novel hum
107	34	64.2	346	3	AAW85739	Membrane-	180	34	64.2	346	6	ADA80588	Human pro
108	34	64.2	346	4	AAW85739	Membrane-	181	34	64.2	346	6	ADA75830	Human pro
109	34	64.2	346	4	AAW85739	Membrane-	182	34	64.2	346	6	ADA38638	Human pro
110	34	64.2	346	4	AAW85739	Membrane-	183	34	64.2	346	6	ADA7055	Human pro
111	34	64.2	346	4	AAW85739	Membrane-	184	34	64.2	346	6	ADA7055	Human pro
112	34	64.2	346	4	AAW85739	Membrane-	185	34	64.2	346	6	ADA35351	Human pro
113	34	64.2	346	5	ABW85811	Human pro	186	34	64.2	346	6	ADA93527	Human pro
114	34	64.2	346	5	ABW85811	Human pro	187	34	64.2	346	6	ADB26877	Human pro
115	34	64.2	346	5	ABW85811	Human pro	188	34	64.2	346	6	ADB31164	Human pro
116	34	64.2	346	5	ABW85811	Human pro	189	34	64.2	346	6	ADA92759	Human pro
117	34	64.2	346	6	ABU58017	Human pro	190	34	64.2	346	6	ADA61092	Homo sapi
118	34	64.2	346	6	ABU58017	Human pro	191	34	64.2	346	6	ADA61092	Homo sapi
119	34	64.2	346	6	ABU58017	Human pro	192	34	64.2	346	6	ADA61092	Homo sapi
120	34	64.2	346	6	ABU58017	Human pro	193	34	64.2	346	6	ADA61092	Homo sapi
121	34	64.2	346	6	ABU58017	Human pro	194	34	64.2	346	6	ADA61092	Homo sapi
122	34	64.2	346	6	ABU58017	Human pro	195	34	64.2	346	6	ADA61092	Homo sapi
123	34	64.2	346	6	ABU58017	Human pro	196	34	64.2	346	6	ADA61092	Homo sapi
124	34	64.2	346	6	ABU58017	Human pro	197	34	64.2	346	6	ADA61092	Homo sapi
125	34	64.2	346	6	ABU58017	Human pro	198	34	64.2	346	6	ADA61092	Homo sapi
126	34	64.2	346	6	ABU58017	Human pro	199	34	64.2	346	6	ADA61092	Homo sapi
127	34	64.2	346	6	ABU58017	Human pro	200	34	64.2	346	6	ADA61092	Homo sapi
128	34	64.2	346	6	ABU58017	Human pro	201	34	64.2	346	6	ADA61092	Homo sapi
129	34	64.2	346	6	ABU58017	Human pro	202	34	64.2	346	6	ADA61092	Homo sapi
130	34	64.2	346	6	ABU58017	Human pro	203	34	64.2	346	6	ADA61092	Homo sapi
131	34	64.2	346	6	ABU58017	Human pro	204	34	64.2	346	6	ADA61092	Homo sapi
132	34	64.2	346	6	ABU58017	Human pro	205	34	64.2	346	6	ADA61092	Homo sapi
133	34	64.2	346	6	ABU58017	Human pro	206	34	64.2	346	6	ADA61092	Homo sapi

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:01:12 ; Search time 55 seconds

(without alignments)  
51.372 Million cell updates/sec

Title: US-10-697-055-7

Perfect score: 53

Sequence: 1 TPRVTGGGAM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

A Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	10	2	AAW48985
2	53	100.0	10	2	AAW09322
3	53	100.0	10	3	AAW12412
4	53	100.0	10	4	AAW05431
5	53	100.0	10	5	AAW10826
6	53	100.0	10	5	AAW66768
7	53	100.0	15	2	AAW10218
8	53	100.0	15	4	AAW86672
9	53	100.0	15	4	AAW86672
10	53	100.0	20	6	ABJ38030
11	53	100.0	551	2	AAW26732
12	53	100.0	551	2	AAW27276
13	53	100.0	560	4	AAW10840
14	53	100.0	561	4	AAW86545
15	53	100.0	631	2	AAW23036
16	48	90.6	10	2	AAW48986
17	48	90.6	10	2	AAW09315
18	48	90.6	10	3	AAW12413
19	48	90.6	10	4	AAW05432
20	48	90.6	10	5	AAW10827
21	46	86.8	15	5	ABG79900
22	44	83.0	11	6	ABJ37950
23	43	81.1	172	6	ABW69745
24	41	77.4	9	6	ABJ37972
25	39	73.6	339	6	ABG72018

26	39	73.6	447	2	AAW29924
27	39	73.6	447	2	AAW58567
28	38	71.7	10	3	AAW01253
29	38	71.7	15	2	AAW86671
30	38	71.7	19	2	AAW47205
31	38	71.7	144	2	AAW47195
32	38	71.7	197	4	ABG16733
33	38	71.7	270	6	ABR41269
34	38	71.7	331	3	AAW50644
35	38	71.7	331	3	AAW13982
36	38	71.7	383	3	AAW13981
37	38	71.7	383	3	AAW50643
38	38	71.7	383	5	AAW72500
39	38	71.7	383	7	ADD30432
40	38	71.7	401	3	AAW13980
41	38	71.7	401	3	AAW50642
42	37	69.8	103	4	AAW59241
43	37	69.8	103	6	ABW55760
44	37	69.8	214	6	ABR01485
45	36	67.9	98	4	ABG02067
46	36	67.9	146	4	ABG15188
47	36	67.9	261	4	ABW65619
48	36	67.9	361	5	ABP65988
49	36	67.9	768	6	ABU17911
50	36	67.9	799	2	AAW19212
51	36	67.9	799	2	AAW56489
52	36	67.9	871	2	AAW70895
53	35	66.0	94	4	AAW10927
54	35	66.0	250	4	AAW18429
55	35	66.0	302	4	AAW18299
56	35	66.0	321	4	AAW70839
57	35	66.0	345	3	AAW93600
58	35	66.0	357	4	AAW04066
59	35	66.0	359	3	ABW23543
60	35	66.0	359	4	ABW04073
61	35	66.0	359	7	ADE59622
62	35	66.0	361	5	AAW10339
63	35	66.0	370	5	AAU77480
64	35	66.0	370	5	AAU77218
65	35	66.0	424	2	AAW97851
66	35	66.0	424	3	AAW71098
67	35	66.0	424	3	AAW84859
68	35	66.0	424	5	AAO22550
69	35	66.0	424	5	AAE18298
70	35	66.0	424	5	AAE16452
71	35	66.0	424	5	ABW09228
72	35	66.0	432	6	ABU11362
73	35	66.0	447	3	AAW31505
74	35	66.0	453	3	AAW31504
75	35	66.0	489	3	AAW36270
76	35	66.0	502	5	ABP73840
77	35	66.0	524	3	AAW31503
78	35	66.0	524	6	ABW39476
79	35	66.0	539	2	AAW33628
80	35	66.0	539	5	ABG93071
81	35	66.0	539	6	ABW33505
82	35	66.0	586	5	ABW33791
83	35	66.0	923	6	ABU36550
84	35	66.0	3238	4	ABW71715
85	34	64.2	86	4	AAW65739
86	34	64.2	86	7	ABW94707
87	34	64.2	88	4	ABG09981
88	34	64.2	96	4	AAW11367
89	34	64.2	128	7	ABW94898
90	34	64.2	142	7	ABW83174
91	34	64.2	162	4	ABG25318
92	34	64.2	169	5	ABG93286
93	34	64.2	194	5	ABG70313
94	34	64.2	206	3	AAW21043
95	34	64.2	206	7	ABW25769
96	34	64.2	236	2	AAW11241
97	34	64.2	247	3	AAW88588
98	34	64.2	260	4	AAW90879

AAW29924	Eimeria a
AAW58567	Eimeria a
AAW01253	Chimaeric
AAW86671	Human cyt
AAW47205	HSV trunc
AAW47195	Herpes si
ABG16733	Novel hum
ABR41269	Human DIT
AAW50644	Arabidops
AAW13982	Arabidops
AAW13981	Arabidops
AAW50643	Arabidops
AAW72500	Arabidops
ADD30432	Plant yie
AAW13980	Arabidops
AAW50642	Arabidops
AAW59241	Propionib
ABW55760	Propionib
ABR01485	Human ant
ABG02067	Novel hum
ABG15188	Novel hum
ABW65619	Drosophil
ABP65988	Bifidobac
ABU17911	Protein e
AAW19212	Corn star
AAW56489	Zea mays
AAW70895	Maize bra
AAW10927	Human pol
AAW18429	Human end
AAW18299	Human end
AAW70839	S. halste
AAW93600	Protein e
AAW04066	Zcytor 10
ABW23543	Murine cy
ABW04073	Zcytor 10
ADE59622	Rat Prote
AAW10339	Novel hum
AAU77480	Murine TS
AAU77218	Mouse thy
AAW97851	Hypersens
AAW71098	Pseudomon
AAW84859	A hyperse
AAO22550	Hypersens
AAE18298	Pseudomon
AAE16452	P. syring
ABW09228	P. syring
ABU11362	Protein e
AAW31505	Arabidops
AAW31504	Arabidops
AAW36270	C. elegan
ABP73840	Candida a
AAW31503	Arabidops
ABW39476	A. thalia
AAW33628	Yeast tra
ABG93071	S. cerevi
ABW33505	Protein s
ABW33791	Herbicida
ABU36550	Protein e
ABW71715	Drosophil
AAW65739	Prohibit
ABW94707	Programme
ABG09981	Novel hum
AAW11367	Human pol
ABW94898	Programme
ABW83174	C4.4A int
ABG25318	Novel hum
ABG93286	C. albica
ABG70313	Human MDD
AAW21043	Human nuc
ABW25769	Human pro
AAW11241	S. pneumo
AAW88588	Death upo
AAW90879	C glutami



Best Local Similarity 77.8%; Pred. No. 6e+02; Mismatches 0; Indels 2; Gaps 0;  
Matches 7; Conservative

Qy 2 PRVTGGGAM 10  
Db 344 PRGPGGGAM 352

RESULT 48  
US-09-895-593-2  
; Sequence 2, Application US/09895593  
; Patent No. US2002016094A1  
; GENERAL INFORMATION:  
; APPLICANT: Pandey, Akhilesh  
; APPLICANT: Ozaki, Katsutoshi  
; APPLICANT: Baumann, Heinz  
; APPLICANT: Levin, Steven D.  
; APPLICANT: Farr, Andrew G.  
; APPLICANT: Ziegler, Steven F.  
; APPLICANT: Leonard, Warren J.  
; APPLICANT: Lodish, Harvey F.  
; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 00-514-E  
; CURRENT APPLICATION NUMBER: US/09/895,593  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/215,658  
; PRIOR FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-895-593-2

Query Match 66.0%; Score 35; DB 9; Length 370;  
Best Local Similarity 77.8%; Pred. No. 6e+02; Mismatches 0; Indels 2; Gaps 0;  
Matches 7; Conservative

Qy 2 PRVTGGGAM 10  
Db 344 PRGPGGGAM 352

RESULT 49  
US-10-425-114-66952  
; Sequence 66952, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)E  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 66952  
; LENGTH: 383  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4756-044-H12\_FLI.pep  
US-10-425-114-66952

Query Match 66.0%; Score 35; DB 12; Length 383;  
Best Local Similarity 85.7%; Pred. No. 6.2e+02; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative

Qy 2 PRVTGGG 8  
Db 87 PRVSGGG 93

RESULT 50  
US-09-835-684-9  
; Sequence 9, Application US/09835684  
; Patent No. US20020019337A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Qiu, Dewen  
; APPLICANT: Remick, Dean  
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE  
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR  
; TITLE OF INVENTION: DESICCATION  
; FILE REFERENCE: 21829/71  
; CURRENT APPLICATION NUMBER: US/09/835,684  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/198,359  
; PRIOR FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
US-09-835-684-9

Query Match 66.0%; Score 35; DB 9; Length 424;  
Best Local Similarity 66.7%; Pred. No. 6.9e+02; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative

Qy 1 TPRVTGGGA 9  
Db 169 TPTATGGGS 177

Search completed: March 18, 2004, 00:59:37  
Job time : 44 secs



; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1701

; LENGTH: 338

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Orthologous to G1652

US-10-374-780A-1701

Query Match

Best Local Similarity 66.0%; Score 35; DB 15; Length 338;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVTGGG 98

RESULT 44

US-09-895-943-3

; Sequence 3, Application US/09895943

; Patent No. US20020068323A1

; GENERAL INFORMATION:

; APPLICANT: Saris, Chris

; APPLICANT: Chang, Ming-Shi

; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

; TITLE OF INVENTION: Uses Thereof

; FILE REFERENCE: 00-514-C

; CURRENT APPLICATION NUMBER: US/09/895,943

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: 60/214,866

; PRIOR FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: TRANSMEM

; LOCATION: (227)..(247)

US-09-895-943-3

Query Match

Best Local Similarity 66.0%; Score 35; DB 9; Length 353;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10

DB 327 PRGPGGGAM 335

RESULT 45

US-09-895-593-3

; Sequence 3, Application US/09895593

; Patent No. US20020160949A1

; GENERAL INFORMATION:

; APPLICANT: Pandey, Akhilesh

; APPLICANT: Ozaki, Katsutoshi

; APPLICANT: Baumann, Heinz

; APPLICANT: Levin, Steven D.

; APPLICANT: Farr, Andrew G.

; APPLICANT: Ziegler, Steven F.

; APPLICANT: Leonard, Warren J.

; APPLICANT: Lodish, Harvey F.

; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

; TITLE OF INVENTION: Uses Thereof

; FILE REFERENCE: 00-514-E

; CURRENT APPLICATION NUMBER: US/09/895,593

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: 60/215,658

; PRIOR FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 353

; TYPE: PRT

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: TRANSMEM

; LOCATION: (227)..(247)

US-09-895-593-3

Query Match

Best Local Similarity 66.0%; Score 35; DB 9; Length 353;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10

DB 327 PRGPGGGAM 335

RESULT 46

US-10-226-872-7

; Sequence 7, Application US/10226872

; Publication No. US20030157118A1

; GENERAL INFORMATION:

; APPLICANT: Cabezon-Silva, Teresa Elisa Virginia

; APPLICANT: Coche, Thierry

; APPLICANT: Gaulis, Swann Romain Jean-Thomas

; APPLICANT: Vinals Y De Bassols, Carlota

; APPLICANT: Casaat, Jean-Pol

; TITLE OF INVENTION: Tumour-Specific Animal Proteins

; FILE REFERENCE: BC45300-1

; CURRENT APPLICATION NUMBER: US/10/226,872

; CURRENT FILING DATE: 2002-12-05

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 361

; TYPE: PRT

; ORGANISM: Human

US-10-226-872-7

Query Match

Best Local Similarity 66.0%; Score 35; DB 14; Length 361;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9

DB 79 PRLAGGGA 86

RESULT 47

US-09-895-943-2

; Sequence 2, Application US/09895943

; Patent No. US20020068323A1

; GENERAL INFORMATION:

; APPLICANT: Saris, Chris

; APPLICANT: Chang, Ming-Shi

; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

; TITLE OF INVENTION: Uses Thereof

; FILE REFERENCE: 00-514-C

; CURRENT APPLICATION NUMBER: US/09/895,943

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: 60/214,866

; PRIOR FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-895-943-2

Query Match

66.0%; Score 35; DB 9; Length 370;

```
Db      21 PRATGGG 27

RESULT 40
US-10-424-599-206428
; Sequence 206428, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206428
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28431C.1.pep
US-10-424-599-206428

Query Match      66.0%; Score 35; DB 12; Length 163;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 PRVTGGG 8
Db      75 PRVSGGG 81

RESULT 41
US-10-424-599-220062
; Sequence 220062, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220062
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_40745C.1.pep
US-10-424-599-220062

Query Match      66.0%; Score 35; DB 12; Length 231;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 TPRVTGGGAM 10
Db      132 TPRITGGGIL 141

RESULT 42
US-10-156-761-9249
; Sequence 9249, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
```

```
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9249
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9249
```

```
Query Match      66.0%; Score 35; DB 14; Length 330;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 PRVTGGGAM 10
Db      152 PRVKGGGAV 160
```

```
RESULT 43
US-10-374-780A-1701
; Sequence 1701, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
```

; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (479)..(479)  
; OTHER INFORMATION: X=any amino acid  
US-10-282-122A-45835

Query Match 67.9%; Score 36; DB 12; Length 768;  
Best Local Similarity 75.0%; Pred. No. 8.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9

DB 374 PRIPGGGA 381

RESULT 36

US-10-369-493-10058  
; Sequence 10058, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 10058  
; LENGTH: 855  
; TYPE: PRT  
; ORGANISM: magnetite-containing magnetic coccus  
US-10-369-493-10058

Query Match 67.9%; Score 36; DB 15; Length 855;  
Best Local Similarity 87.5%; Pred. No. 9.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9

DB 72 PRVHGGGA 79

RESULT 37

US-10-424-599-207115  
; Sequence 207115, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 207115  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(153)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29050C.1.pep  
US-10-424-599-207115

Query Match 66.0%; Score 35; DB 12; Length 69;

Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 33 SPEVTGGG 40

RESULT 38

US-10-424-599-258345  
; Sequence 258345, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 258345  
; LENGTH: 94  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_7530C.1.pep  
US-10-424-599-258345

Query Match 66.0%; Score 35; DB 12; Length 94;  
Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8

DB 48 PRATGGG 54

RESULT 39

US-10-424-599-255296  
; Sequence 255296, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 255296  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(153)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_72554C.1.pep  
US-10-424-599-255296

Query Match 66.0%; Score 35; DB 12; Length 153;  
Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8

DB 48 PRATGGG 54

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70152
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73061A02_FLI.pep
US-10-425-114-70152

Query Match 67.9%; Score 36; DB 12; Length 560;
Best Local Similarity 77.8%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 205 TPRSTGSGA 213

RESULT 34
US-10-369-493-20821
; Sequence 20821, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20821
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20821

Query Match 67.9%; Score 36; DB 15; Length 600;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VTGGGAM 10
Db 22 VTGGGAM 28

RESULT 35
US-10-282-122A-45835
; Sequence 45835, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELIPIA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45835
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (114)..(114)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (116)..(116)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (319)..(319)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (348)..(348)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (349)..(349)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (360)..(360)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (371)..(371)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (382)..(382)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (383)..(383)
```

```
Query Match      67.9%; Score 36; DB 14; Length 244;
Best Local Similarity 70.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
DB 74 TPRPTGGAKM 83

RESULT 29
US-10-369-493-15690
; Sequence 15690, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15690
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15690

Query Match      67.9%; Score 36; DB 15; Length 276;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
DB 228 TPRLTGSGS 236

RESULT 30
US-10-369-493-16082
; Sequence 16082, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16082
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16082

Query Match      67.9%; Score 36; DB 15; Length 276;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
DB 228 TPRLTGSGS 236

RESULT 31
US-10-369-493-15324
; Sequence 15324, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15324
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15324

Query Match      67.9%; Score 36; DB 15; Length 294;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
DB 242 TPRLTGSGS 250

RESULT 32
US-10-425-114-64171
; Sequence 64171, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64171
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-009-E11_FLI.pgp
US-10-425-114-64171

Query Match      67.9%; Score 36; DB 12; Length 346;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
DB 54 TPRVAGGG 61

RESULT 33
US-10-425-114-70152
; Sequence 70152, Application US/10425114
; Publication No. US20040034888A1
```

; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-12713

Query Match 69.8%; Score 37; DB 14; Length 748;  
Best Local Similarity 75.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9  
||:||||  
Db 593 PRINGGA 600

## RESULT 25

US-10-424-599-180013  
; Sequence 180013, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 180013  
; LENGTH: 39  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_133567C.1.pap  
US-10-424-599-180013

Query Match 67.9%; Score 36; DB 12; Length 39;  
Best Local Similarity 75.0%; Pred. No. 44;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9  
||:||||  
Db 3 PKVSGGA 10

## RESULT 26

US-10-156-761-13012  
; Sequence 13012, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13012  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13012

Query Match 67.9%; Score 36; DB 14; Length 54;

Best Local Similarity 60.0%; Pred. No. 60;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 10  
||:||||  
Db 28 SPYMTGGGA 37

## RESULT 27

US-10-029-386-27829  
; Sequence 27829, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 27829  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL33476.4  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4  
; OTHER INFORMATION: SWISSPROT HIT: O08788, EVALUE 1.60e+00  
US-10-029-386-27829

Query Match 67.9%; Score 36; DB 14; Length 145;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
||:||||  
Db 25 TPRATGGA 33

## RESULT 28

US-10-156-761-15056  
; Sequence 15056, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 15056  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-15056

```
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2854
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2373
US-10-374-780A-2854

Query Match          71.7%; Score 38; DB 15; Length 383;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8
Db 62 TPRVTGGG 69

RESULT 21
US-10-424-599-232152
; Sequence 232152, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232152
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51655C.1.pep
US-10-424-599-232152

Query Match          69.8%; Score 37; DB 12; Length 70;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8
Db 20 TPRNTGGG 27

RESULT 22
US-10-424-599-241541
; Sequence 241541, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```

```
; SEQ ID NO 241541
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60138C.1.pep
US-10-424-599-241541

Query Match          69.8%; Score 37; DB 12; Length 121;
Best Local Similarity 77.8%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 9
Db 55 TPTFTGGG 63

RESULT 23
US-10-306-762-109
; Sequence 109, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; TITLE OF INVENTION: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 313
; TYPE: PRT
; ORGANISM: C. aurantiacus (22971932)
US-10-306-762-109

Query Match          69.8%; Score 37; DB 14; Length 313;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGG 8
Db 98 SPKVTGGG 105

RESULT 24
US-10-156-761-12713
; Sequence 12713, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12713
; LENGTH: 748
```

```

; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghassemlan, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyuki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCE: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 124
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-259-194A-124

Query Match          73.6%; Score 39; DB 15; Length 665;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
Db 27 TPRUSGGGTL 36

RESULT 18
; Sequence 269734, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 269734
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: uncore
; LOCATION: (1)..(115)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; . OTHER INFORMATION: Clone ID: PAT_MRT3847_8558C.1.pap
US-10-424-599-269734

Query Match          71.7%; Score 38; DB 12; Length 115;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9

```

```

Db 7 TPRTTGGGA 15

RESULT 19
; US-10-425-114-70467
; Sequence 70467, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70467
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMOL17259F06_FLI.pap
US-10-425-114-70467

Query Match          71.7%; Score 38; DB 12; Length 210;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
Db 73 PRVAGGA 80

RESULT 20
; US-10-374-780A-2854
; Sequence 2854, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468

```



```

; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-051-874-86

Query Match      73.6%; Score 39; DB 15; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 PRVTGGGA 9
DB      43 PRITGGGS 50

RESULT 16
US-10-051-874-87
; Sequence 87, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G

```

```

; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol BA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-051-874-87

Query Match      73.6%; Score 39; DB 15; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 PRVTGGGA 9
DB      43 PRITGGGS 50

RESULT 17
US-10-259-194A-124
; Sequence 124, Application US/10259194A

```

APPLICANT:	ALSOBROOK, LI, JOHN
APPLICANT:	COLMAN, STEVEN D
APPLICANT:	SPYTEK, KIMBERLY A
APPLICANT:	BOLDIG, FERENC
APPLICANT:	VERNET, CORINE AM
APPLICANT:	LI, LI
APPLICANT:	SHENOY, SURESH G
APPLICANT:	CASMAN, STACIE J
APPLICANT:	GUO, XIAOJIA SASHA
APPLICANT:	EDINGER, SHLOMIT R
APPLICANT:	MACDOUGALL, JOHN R
APPLICANT:	MALYANKAR, URIEL M
APPLICANT:	PATTURAJAN, MEERA
APPLICANT:	SHIMKETS, RICHARD A
APPLICANT:	PENA, CAROL RA
APPLICANT:	TCHERNEV, VELIZAR T
APPLICANT:	ZERHUSEN, BRYAN D

```

; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennnda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 153
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-042-865-153

Query Match      73.6%; Score 39; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
Db      43 PRITGGGS 50

RESULT 11
US-10-109-616-2
; Sequence 2, Application US/10109616
; Publication No. US20030167484A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING CHANNEL
; TITLE OF INVENTION: ACTIVATING PROTEASE 1 (CAP1) GENE DISRUPTIONS
; FILE REFERENCE: R-490
; CURRENT APPLICATION NUMBER: US/10/109,616
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/280,509
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/311,055
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-109-616-2

Query Match      73.6%; Score 39; DB 14; Length 339;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
Db      43 PRITGGGS 50

```

```

; Sequence 151, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennnda
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-042-865-151

Query Match      73.6%; Score 39; DB 12; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 PRVTGGGA 9
Db      43 PRITGGGS 50

RESULT 13
US-10-042-865-152
; Sequence 152, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J

```

; CURRENT FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: US/09/692,170C  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 09/534,639  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/075,257  
; PRIOR FILING DATE: 1998-05-11  
; PRIOR APPLICATION NUMBER: US 09/021,298  
; PRIOR FILING DATE: 1998-02-10  
; PRIOR APPLICATION NUMBER: US 08/950,064  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: US 08/747,488  
; PRIOR FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Variant human cytomegalovirus peptide epitope  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (10)..(10)  
; OTHER INFORMATION: Xaa = L, F or M  
; US-10-405-231-8

Query Match 90.6%; Score 48; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0

Qy 1 TPRTVGGGA 9  
| | | | |  
Db 1 TPRTVGGGA 9

RESULT 9  
US-10-051-874-89  
; Sequence 89, Application US/10051874  
; Publication No. US20040005557A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Coleman, Steven D  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Vernet, Corine AM  
; APPLICANT: Li, Li  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Guo, Xiaojia Saaha  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: MacDougall, John R  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Pena, Carol EA  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Miller, Charles E  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Baumgartner, Jason C  
; APPLICANT: Herrman, John L  
; APPLICANT: Peyman, John A  
; APPLICANT: Gorman, Linda  
; APPLICANT: Mezes, Peter D  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Grosse, William M  
; APPLICANT: Liu, Xiaohong

; APPLICANT: Ellerman, Karen  
; APPLICANT: Rotherberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Burgess, Catherine E  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; TITLE OF INVENTION: USING THE SAME  
; FILE REFERENCE: 21402-245  
; CURRENT APPLICATION NUMBER: US/10/051,874  
; CURRENT FILING DATE: 2002-09-25  
; PRIOR APPLICATION NUMBER: 60/268,595  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/325,306  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/262,567  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,409  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/262,454  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/276,777  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/291,672  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/330,336  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/265,530  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/261,376  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 269  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 89  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-10-051-874-89

Query Match 73.6%; Score 39; DB 15; Length 285;  
Best Local Similarity 75.0%; Pred. No. 99;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
| | | | |  
Db 43 PRITGGGS 50

RESULT 10  
US-10-042-865-153  
; Sequence 153, Application US/10042865  
; Publication No. US20040029216A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine A.M  
; APPLICANT: Taylor, Sarah  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Miller, Charles E  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Boldog, Ference L  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Rotherberg, Mark E  
; APPLICANT: Ellerman, Karen



```
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human cytomegalovirus
US-10-238-607-7

Query Match          100.0%; Score 53; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10
Db 1 TPRVTGGGAM 10

RESULT 2
US-10-405-231-7
; Sequence 7, Application US/10405231
; Publication No. US20030190328A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-346
; CURRENT APPLICATION NUMBER: US/10/405,231
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US/09/692,170C
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US/09/534,639
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US/09/075,257
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: US/09/021,298
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US/08/950,064
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: US/08/747,488
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human cytomegalovirus
US-10-405-231-7

Query Match          100.0%; Score 53; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10
Db 1 TPRVTGGGAM 10

RESULT 3
US-09-812-079A-2
; Sequence 2, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS COMPOUNDS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15

Query Match          100.0%; Score 53; DB 14; Length 561;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10
Db 417 TPRVTGGGAM 426

RESULT 4
US-10-434-982-2
; Sequence 2, Application US/10434982
; Publication No. US20030199673A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS COMPOUNDS
; FILE REFERENCE: 5018C
; CURRENT APPLICATION NUMBER: US/10/434,982
; CURRENT FILING DATE: 2003-05-10
; PRIOR APPLICATION NUMBER: US/09/812,079
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: US/60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US/60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-434-982-2

Query Match          100.0%; Score 53; DB 14; Length 561;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPRVTGGGAM 10
Db 417 TPRVTGGGAM 426

RESULT 5
US-10-223-538-6
; Sequence 6, Application US/10223538
; Publication No. US20030120060A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Berencsi, Klara
; Kari, Ceaba
; TITLE OF INVENTION: No. US20030120060A1el Cytomegalovirus DNA Constructs and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

381	34	64.2	346	14	US-10-145-630-432	Sequence 432, App
382	34	64.2	346	14	US-10-145-747-432	Sequence 432, App
383	34	64.2	346	14	US-10-145-752-432	Sequence 432, App
384	34	64.2	346	14	US-10-145-754-432	Sequence 432, App
385	34	64.2	346	14	US-10-145-755-432	Sequence 432, App
386	34	64.2	346	14	US-10-145-818-432	Sequence 432, App
387	34	64.2	346	14	US-10-145-820-432	Sequence 432, App
388	34	64.2	346	14	US-10-145-872-432	Sequence 432, App
389	34	64.2	346	14	US-10-145-873-432	Sequence 432, App
390	34	64.2	346	14	US-10-147-481-432	Sequence 432, App
391	34	64.2	346	14	US-10-147-482-432	Sequence 432, App
392	34	64.2	346	14	US-10-147-503-432	Sequence 432, App
393	34	64.2	346	14	US-10-147-522-432	Sequence 432, App
394	34	64.2	346	14	US-10-152-401-432	Sequence 432, App
395	34	64.2	346	14	US-10-157-783-432	Sequence 432, App
396	34	64.2	346	14	US-10-158-792-432	Sequence 432, App
397	34	64.2	346	14	US-10-158-822-432	Sequence 432, App
398	34	64.2	346	14	US-10-143-035-432	Sequence 432, App
399	34	64.2	346	14	US-10-145-751-432	Sequence 432, App
400	34	64.2	346	14	US-10-145-822-432	Sequence 432, App
401	34	64.2	346	14	US-10-145-824-432	Sequence 432, App
402	34	64.2	346	14	US-10-145-827-432	Sequence 432, App
403	34	64.2	346	14	US-10-145-859-432	Sequence 432, App
404	34	64.2	346	14	US-10-145-875-432	Sequence 432, App
405	34	64.2	346	14	US-10-145-877-432	Sequence 432, App
406	34	64.2	346	14	US-10-145-958-432	Sequence 432, App
407	34	64.2	346	14	US-10-146-787-432	Sequence 432, App
408	34	64.2	346	14	US-10-146-790-432	Sequence 432, App
409	34	64.2	346	14	US-10-146-793-432	Sequence 432, App
410	34	64.2	346	14	US-10-147-480-432	Sequence 432, App
411	34	64.2	346	14	US-10-147-485-432	Sequence 432, App
412	34	64.2	346	14	US-10-147-486-432	Sequence 432, App
413	34	64.2	346	14	US-10-147-487-432	Sequence 432, App
414	34	64.2	346	14	US-10-147-490-432	Sequence 432, App
415	34	64.2	346	14	US-10-147-494-432	Sequence 432, App
416	34	64.2	346	14	US-10-147-498-432	Sequence 432, App
417	34	64.2	346	14	US-10-147-514-432	Sequence 432, App
418	34	64.2	346	14	US-10-147-524-432	Sequence 432, App
419	34	64.2	346	14	US-10-152-379-432	Sequence 432, App
420	34	64.2	346	14	US-10-152-394-432	Sequence 432, App
421	34	64.2	346	14	US-10-152-406-432	Sequence 432, App
422	34	64.2	346	14	US-10-156-847-432	Sequence 432, App
423	34	64.2	346	14	US-10-157-778-432	Sequence 432, App
424	34	64.2	346	14	US-10-157-799-432	Sequence 432, App
425	34	64.2	346	14	US-10-160-504-432	Sequence 432, App
426	34	64.2	346	14	US-10-145-634-432	Sequence 432, App
427	34	64.2	346	14	US-10-147-520-432	Sequence 432, App
428	34	64.2	346	14	US-10-157-781-432	Sequence 432, App
429	34	64.2	346	14	US-10-176-989-432	Sequence 432, App
430	34	64.2	346	14	US-10-147-491-432	Sequence 432, App
431	34	64.2	346	14	US-10-152-378-432	Sequence 432, App
432	34	64.2	346	14	US-10-152-382-432	Sequence 432, App
433	34	64.2	346	14	US-10-152-383-432	Sequence 432, App
434	34	64.2	346	14	US-10-152-384-432	Sequence 432, App
435	34	64.2	346	14	US-10-152-387-432	Sequence 432, App
436	34	64.2	346	14	US-10-152-389-432	Sequence 432, App
437	34	64.2	346	14	US-10-152-390-432	Sequence 432, App
438	34	64.2	346	14	US-10-152-392-432	Sequence 432, App
439	34	64.2	346	14	US-10-153-756-432	Sequence 432, App
440	34	64.2	346	14	US-10-157-784-432	Sequence 432, App
441	34	64.2	346	14	US-10-157-797-432	Sequence 432, App
442	34	64.2	346	14	US-10-158-491-432	Sequence 432, App
443	34	64.2	346	14	US-10-142-762-432	Sequence 432, App
444	34	64.2	346	14	US-10-142-763-432	Sequence 432, App
445	34	64.2	346	14	US-10-142-766-432	Sequence 432, App
446	34	64.2	346	14	US-10-145-625-432	Sequence 432, App
447	34	64.2	346	14	US-10-145-627-432	Sequence 432, App
448	34	64.2	346	14	US-10-145-960-432	Sequence 432, App
449	34	64.2	346	14	US-10-145-962-432	Sequence 432, App
450	34	64.2	346	14	US-10-146-789-432	Sequence 432, App
451	34	64.2	346	14	US-10-147-483-432	Sequence 432, App
452	34	64.2	346	14	US-10-147-496-432	Sequence 432, App
453	34	64.2	346	14	US-10-147-505-432	Sequence 432, App
454	34	64.2	346	14	US-10-147-516-432	Sequence 432, App
455	34	64.2	346	14	US-10-152-398-432	Sequence 432, App
456	34	64.2	346	14	US-10-139-980-432	Sequence 432, App
457	34	64.2	346	14	US-10-145-750-432	Sequence 432, App
458	34	64.2	346	14	US-10-152-373-432	Sequence 432, App
459	34	64.2	346	14	US-10-223-081-130	Sequence 130, App
460	34	64.2	346	14	US-10-121-044-432	Sequence 432, App
461	34	64.2	346	14	US-10-121-055-432	Sequence 432, App
462	34	64.2	346	14	US-10-121-057-432	Sequence 432, App
463	34	64.2	346	14	US-10-121-058-432	Sequence 432, App
464	34	64.2	346	14	US-10-121-059-432	Sequence 432, App
465	34	64.2	346	14	US-10-121-060-432	Sequence 432, App
466	34	64.2	346	14	US-10-123-109-432	Sequence 432, App
467	34	64.2	346	14	US-10-123-154-432	Sequence 432, App
468	34	64.2	346	14	US-10-123-157-432	Sequence 432, App
469	34	64.2	346	14	US-10-123-906-432	Sequence 432, App
470	34	64.2	346	14	US-10-124-814-432	Sequence 432, App
471	34	64.2	346	14	US-10-124-816-432	Sequence 432, App
472	34	64.2	346	14	US-10-124-820-432	Sequence 432, App
473	34	64.2	346	14	US-10-125-704-432	Sequence 432, App
474	34	64.2	346	14	US-10-125-927-432	Sequence 432, App
475	34	64.2	346	14	US-10-223-082-130	Sequence 130, App
476	34	64.2	346	14	US-10-142-889-432	Sequence 432, App
477	34	64.2	346	14	US-10-145-874-432	Sequence 432, App
478	34	64.2	346	14	US-10-147-497-432	Sequence 432, App
479	34	64.2	346	14	US-10-152-371-432	Sequence 432, App
480	34	64.2	346	14	US-10-152-374-432	Sequence 432, App
481	34	64.2	346	14	US-10-152-375-432	Sequence 432, App
482	34	64.2	346	14	US-10-152-377-432	Sequence 432, App
483	34	64.2	346	14	US-10-152-386-432	Sequence 432, App
484	34	64.2	346	14	US-10-152-391-432	Sequence 432, App
485	34	64.2	346	14	US-10-152-399-432	Sequence 432, App
486	34	64.2	346	14	US-10-156-848-432	Sequence 432, App
487	34	64.2	346	14	US-10-157-785-432	Sequence 432, App
488	34	64.2	346	14	US-10-157-794-432	Sequence 432, App
489	34	64.2	346	14	US-10-157-796-432	Sequence 432, App
490	34	64.2	346	14	US-10-160-500-432	Sequence 432, App
491	34	64.2	346	14	US-10-121-046-432	Sequence 432, App
492	34	64.2	346	14	US-10-123-156-432	Sequence 432, App
493	34	64.2	346	14	US-10-123-214-432	Sequence 432, App
494	34	64.2	346	14	US-10-125-805-432	Sequence 432, App
495	34	64.2	346	14	US-10-124-821-432	Sequence 432, App
496	34	64.2	346	14	US-10-152-385-432	Sequence 432, App
497	34	64.2	346	14	US-10-152-393-432	Sequence 432, App
498	34	64.2	346	14	US-10-152-396-432	Sequence 432, App
499	34	64.2	346	14	US-10-153-552-432	Sequence 432, App
500	34	64.2	346	14	US-10-153-840-432	Sequence 432, App

## ALIGNMENTS

## RESULT 1

US-10-238-607-7  
; Sequence 7, Application US/10238607  
; Publication No. US20030118602A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: HCMV- REACTIVE T CELLS AND USES THEREFOR  
; FILE REFERENCE: 1954-398  
; CURRENT APPLICATION NUMBER: US/10/238,607  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 09/692,170  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 09/534,639  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/075,257  
; PRIOR FILING DATE: 1998-05-11  
; PRIOR APPLICATION NUMBER: US 09/021,298  
; PRIOR FILING DATE: 1998-02-10  
; PRIOR APPLICATION NUMBER: US 08/950,064  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: US 08/747,488





89	34	64.2	346	9	US-09-992-598-197	Sequence 197, App	162	34	64.2	346	14	US-10-143-114-432	Sequence 432, App
90	34	64.2	346	9	US-09-989-293A-197	Sequence 197, App	163	34	64.2	346	14	US-10-140-002-432	Sequence 432, App
91	34	64.2	346	9	US-09-989-735-197	Sequence 197, App	164	34	64.2	346	14	US-10-142-419-432	Sequence 432, App
92	34	64.2	346	9	US-09-990-444-197	Sequence 197, App	165	34	64.2	346	14	US-10-123-262-432	Sequence 432, App
93	34	64.2	346	9	US-09-991-181-197	Sequence 197, App	166	34	64.2	346	14	US-10-142-423-432	Sequence 432, App
94	34	64.2	346	9	US-09-989-730-197	Sequence 197, App	167	34	64.2	346	14	US-10-121-050-432	Sequence 432, App
95	34	64.2	346	9	US-09-990-436-197	Sequence 197, App	168	34	64.2	346	14	US-10-141-755-432	Sequence 432, App
96	34	64.2	346	9	US-09-993-687-197	Sequence 197, App	169	34	64.2	346	14	US-10-143-032-432	Sequence 432, App
97	34	64.2	346	10	US-09-989-734-197	Sequence 197, App	170	34	64.2	346	14	US-10-123-108-432	Sequence 432, App
98	34	64.2	346	10	US-09-997-653-197	Sequence 197, App	171	34	64.2	346	14	US-10-123-236-432	Sequence 432, App
99	34	64.2	346	10	US-09-993-628-197	Sequence 197, App	172	34	64.2	346	14	US-10-123-261-432	Sequence 432, App
100	34	64.2	346	10	US-09-997-628-197	Sequence 197, App	173	34	64.2	346	14	US-10-140-921-432	Sequence 432, App
101	34	64.2	346	10	US-09-997-666-197	Sequence 197, App	174	34	64.2	346	14	US-10-140-928-432	Sequence 432, App
102	34	64.2	346	10	US-09-990-438-197	Sequence 197, App	175	34	64.2	346	14	US-10-121-045-432	Sequence 432, App
103	34	64.2	346	10	US-09-990-562-197	Sequence 197, App	176	34	64.2	346	14	US-10-123-292-432	Sequence 432, App
104	34	64.2	346	10	US-09-990-711-197	Sequence 197, App	177	34	64.2	346	14	US-10-123-903-432	Sequence 432, App
105	34	64.2	346	10	US-09-989-726-197	Sequence 197, App	178	34	64.2	346	14	US-10-124-819-432	Sequence 432, App
106	34	64.2	346	10	US-09-988-156-197	Sequence 197, App	179	34	64.2	346	14	US-10-124-822-432	Sequence 432, App
107	34	64.2	346	10	US-09-990-437-197	Sequence 197, App	180	34	64.2	346	14	US-10-140-925-432	Sequence 432, App
108	34	64.2	346	10	US-09-991-157-197	Sequence 197, App	181	34	64.2	346	14	US-10-160-498-432	Sequence 432, App
109	34	64.2	346	10	US-09-997-514-197	Sequence 197, App	182	34	64.2	346	14	US-10-124-824-432	Sequence 432, App
110	34	64.2	346	10	US-09-997-573-197	Sequence 197, App	183	34	64.2	346	14	US-10-127-825A-432	Sequence 432, App
111	34	64.2	346	10	US-09-991-172-197	Sequence 197, App	184	34	64.2	346	14	US-10-127-829A-432	Sequence 432, App
112	34	64.2	346	10	US-09-997-683-197	Sequence 197, App	185	34	64.2	346	14	US-10-127-835A-432	Sequence 432, App
113	34	64.2	346	10	US-09-997-559-197	Sequence 197, App	186	34	64.2	346	14	US-10-127-839A-432	Sequence 432, App
114	34	64.2	346	10	US-09-997-601-197	Sequence 197, App	187	34	64.2	346	14	US-10-127-901A-432	Sequence 432, App
115	34	64.2	346	10	US-09-990-443-197	Sequence 197, App	188	34	64.2	346	14	US-10-128-693A-432	Sequence 432, App
116	34	64.2	346	10	US-09-991-854-197	Sequence 197, App	189	34	64.2	346	14	US-10-131-813A-432	Sequence 432, App
117	34	64.2	346	10	US-09-997-628-197	Sequence 197, App	190	34	64.2	346	14	US-10-131-818A-432	Sequence 432, App
118	34	64.2	346	10	US-09-997-683-197	Sequence 197, App	191	34	64.2	346	14	US-10-131-823A-432	Sequence 432, App
119	34	64.2	346	10	US-09-989-729A-197	Sequence 197, App	192	34	64.2	346	14	US-10-131-824A-432	Sequence 432, App
120	34	64.2	346	10	US-09-997-349-197	Sequence 197, App	193	34	64.2	346	14	US-10-131-830A-432	Sequence 432, App
121	34	64.2	346	10	US-09-997-440-197	Sequence 197, App	194	34	64.2	346	14	US-10-131-837A-432	Sequence 432, App
122	34	64.2	346	10	US-09-990-440-197	Sequence 197, App	195	34	64.2	346	14	US-10-137-872A-432	Sequence 432, App
123	34	64.2	346	10	US-09-993-459-197	Sequence 197, App	196	34	64.2	346	14	US-10-147-500-432	Sequence 432, App
124	34	64.2	346	10	US-09-997-542-197	Sequence 197, App	197	34	64.2	346	14	US-10-147-502-432	Sequence 432, App
125	34	64.2	346	10	US-09-993-748-197	Sequence 197, App	198	34	64.2	346	14	US-10-147-515-432	Sequence 432, App
126	34	64.2	346	10	US-09-990-439-197	Sequence 197, App	199	34	64.2	346	14	US-10-147-517-432	Sequence 432, App
127	34	64.2	346	10	US-09-990-428-197	Sequence 197, App	200	34	64.2	346	14	US-10-147-526-432	Sequence 432, App
128	34	64.2	346	10	US-09-989-328-197	Sequence 197, App	201	34	64.2	346	14	US-10-147-527-432	Sequence 432, App
129	34	64.2	346	10	US-09-993-582-197	Sequence 197, App	202	34	64.2	346	14	US-10-121-041-432	Sequence 432, App
130	34	64.2	346	10	US-09-941-593-197	Sequence 197, App	203	34	64.2	346	14	US-10-121-043-432	Sequence 432, App
131	34	64.2	346	10	US-09-992-521-197	Sequence 197, App	204	34	64.2	346	14	US-10-121-047-432	Sequence 432, App
132	34	64.2	346	10	US-09-997-333-197	Sequence 197, App	205	34	64.2	346	14	US-10-123-215-432	Sequence 432, App
133	34	64.2	346	10	US-09-997-384-197	Sequence 197, App	206	34	64.2	346	14	US-10-123-902-432	Sequence 432, App
134	34	64.2	346	10	US-09-998-041-197	Sequence 197, App	207	34	64.2	346	14	US-10-123-908-432	Sequence 432, App
135	34	64.2	346	10	US-09-997-585-197	Sequence 197, App	208	34	64.2	346	14	US-10-123-909-432	Sequence 432, App
136	34	64.2	346	10	US-09-997-614-197	Sequence 197, App	209	34	64.2	346	14	US-10-123-910-432	Sequence 432, App
137	34	64.2	346	10	US-09-833-041-77	Sequence 77, Appl	210	34	64.2	346	14	US-10-124-813-432	Sequence 432, App
138	34	64.2	346	10	US-09-989-862-197	Sequence 197, App	211	34	64.2	346	14	US-10-124-817-432	Sequence 432, App
139	34	64.2	346	10	US-09-997-529-197	Sequence 197, App	212	34	64.2	346	14	US-10-125-922-432	Sequence 432, App
140	34	64.2	346	10	US-09-989-725-197	Sequence 197, App	213	34	64.2	346	14	US-10-125-924-432	Sequence 432, App
141	34	64.2	346	11	US-09-989-733-197	Sequence 197, App	214	34	64.2	346	14	US-10-140-860-432	Sequence 432, App
142	34	64.2	346	11	US-09-992-643-197	Sequence 197, App	215	34	64.2	346	14	US-10-142-417-432	Sequence 432, App
143	34	64.2	346	11	US-09-833-245-2068	Sequence 2068, App	216	34	64.2	346	14	US-10-147-519-432	Sequence 432, App
144	34	64.2	346	12	US-10-147-493-432	Sequence 432, App	217	34	64.2	346	14	US-10-157-782-432	Sequence 432, App
145	34	64.2	346	12	US-10-145-127-432	Sequence 432, App	218	34	64.2	346	14	US-10-152-395-432	Sequence 432, App
146	34	64.2	346	12	US-10-160-503-432	Sequence 432, App	219	34	64.2	346	14	US-10-125-926A-432	Sequence 432, App
147	34	64.2	346	12	US-10-143-118-432	Sequence 432, App	220	34	64.2	346	14	US-10-125-930A-432	Sequence 432, App
148	34	64.2	346	12	US-10-144-993-432	Sequence 432, App	221	34	64.2	346	14	US-10-127-831A-432	Sequence 432, App
149	34	64.2	346	12	US-10-158-787-432	Sequence 432, App	222	34	64.2	346	14	US-10-127-837A-432	Sequence 432, App
150	34	64.2	346	12	US-10-081-056-130	Sequence 130, App	223	34	64.2	346	14	US-10-127-838B-432	Sequence 432, App
151	34	64.2	346	13	US-10-001-054-16	Sequence 16, Appl	224	34	64.2	346	14	US-10-127-842A-432	Sequence 432, App
152	34	64.2	346	14	US-10-028-072-432	Sequence 432, App	225	34	64.2	346	14	US-10-127-843A-432	Sequence 432, App
153	34	64.2	346	14	US-10-121-049-432	Sequence 432, App	226	34	64.2	346	14	US-10-127-845A-432	Sequence 432, App
154	34	64.2	346	14	US-10-123-904-432	Sequence 432, App	227	34	64.2	346	14	US-10-127-846A-432	Sequence 432, App
155	34	64.2	346	14	US-10-140-470-432	Sequence 432, App	228	34	64.2	346	14	US-10-127-848A-432	Sequence 432, App
156	34	64.2	346	14	US-10-175-746-432	Sequence 432, App	229	34	64.2	346	14	US-10-127-849A-432	Sequence 432, App
157	34	64.2	346	14	US-10-176-918-432	Sequence 432, App	230	34	64.2	346	14	US-10-127-850A-432	Sequence 432, App
158	34	64.2	346	14	US-10-176-921-432	Sequence 432, App	231	34	64.2	346	14	US-10-127-851A-432	Sequence 432, App
159	34	64.2	346	14	US-10-137-865-432	Sequence 432, App	232	34	64.2	346	14	US-10-128-694A-432	Sequence 432, App
160	34	64.2	346	14	US-10-140-474-432	Sequence 432, App	233	34	64.2	346	14	US-10-128-686A-432	Sequence 432, App
161	34	64.2	346	14	US-10-142-431-432	Sequence 432, App	234	34	64.2	346	14	US-10-128-690A-432	Sequence 432, App

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 00:55:18 ; Search time 39 Seconds  
(without alignments)

66.009 Million cell updates/sec  
(without arguments)

Title: US-10-697-055-7

Perfect score: 53

Sequence: 1 TPRVTGGGAM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 104540 seqs, 257433775 residues

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seq	length: 20000000
----------------	------------------

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

- 1: /cgn2\_6/ptodata/1/pubpaa/US07 PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06 PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/ECTUS PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08 PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B PUBCOMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C PUBCOMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pcp.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pcp.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60 PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	53	100.0	10	14	US-10-238-607-7	Sequence 7, Appli
2	53	100.0	10	14	US-10-405-231-7	Sequence 7, Appli
3	53	100.0	561	9	US-09-813-079A-2	Sequence 2, Appli
4	53	100.0	4	561	US-10-434-982-2	Sequence 2, Appli
5	53	100.0	579	14	US-10-223-538-6	Sequence 6, Appli
6	53	100.0	579	14	US-10-223-538-8	Sequence 8, Appli
7	48	90.6	10	14	US-10-238-607-8	Sequence 8, Appli
8	48	90.6	10	14	US-10-405-231-8	Sequence 8, Appli
9	39	73.6	285	15	US-10-051-874-89	Sequence 89, Appli
10	39	73.6	339	12	US-10-042-865-153	Sequence 153, App
11	39	73.6	339	14	US-10-109-616-2	Sequence 2, Appli
12	39	73.6	342	12	US-10-042-865-151	Sequence 151, App
13	39	73.6	342	12	US-10-042-865-152	Sequence 152, App
14	39	73.6	342	12	US-10-042-865-154	Sequence 154, App
15	39	73.6	342	15	US-10-051-874-86	Sequence 86, Appli

Db 40 TAITGGGA 48

## RESULT 49

JQ1252

hypothetical 16.7K protein - chrysanthemum virus B

C;Species: chrysanthemum virus B

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 30-Sep-1993

C;Accession: JQ1252

B;Levay, K.; Zavrilev, S.

J. Gen. Virol. 72, 2333-2337, 1991

A;Title: Nucleotide sequence and gene organization of the 3'-terminal region of chrysanthemum virus B

A;Reference number: JQ1246; MUID:92013948; PMID:1919520

A;Accession: JQ1252

A;Molecule type: genomic RNA

A;Residues: 1-164 &lt;LEV&gt;

Query Match 62.3%; Score 33; DB 2; Length 164;

Best Local Similarity 62.5%; Pred. No. 99;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

: ||: |||

Db 54 SPRILGGG 61

## RESULT 50

AD3294

hypothetical protein BMEI0338 [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002

C;Accession: AD3294

R;DelVecchio, V.G.; Kaputrai, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, A.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AD3294

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-180 &lt;KUR&gt;

A;Cross-references: GB:AE008917; PIDN:AAU51519.1; PID:gl7982235; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI0338

A;Map position: I

Query Match 62.3%; Score 33; DB 2; Length 180;

Best Local Similarity 75.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

: ||: |||

Db 127 TPEVTSGG 134

Search completed: March 7, 2004, 13:02:55

Job time : 36 secs

```

Db          ||| |||
          97 TPRVTGGG 104

RESULT 44
I38740
I: kappa chain V region (Py42) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 21-Jan-2000
C:Accession: I38740
R:Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.
J. Biol. Chem. 266, 6607-6613, 1991
A:Title: Heavy and light chain variable region sequences and antibody properties of anti
A:Reference number: A38740; MUID:91177923; PMID:1706720
A:Accession: I38740
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-111 <RUF>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:19-93/Domain: immunoglobulin homology <IMM>

Query Match 62.3%; Score 33; DB 2; Length 111;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
||| |||
Db 97 TPRVTGGG 104

RESULT 45
JC2270
PL7-6 antibody light chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 21-Jan-2000
C:Accession: JC2270; PC2187
R:Kurone, T.; Katayama, M.; Murakami, K.; Hashino, K.; Kamihagi, K.; Yasumoto, M.; Kato,
J. Biochem. 115, 608-614, 1994
A:Title: Expression of recombinant mouse/human chimeric antibody specific to human GMP-1
A:Reference number: JC2269; MUID:94334310; PMID:7520038
A:Accession: JC2270
A:Molecule type: mRNA
A:Residues: 1-113 <KUR>
A:Accession: PC2187
A:Molecule type: protein
A:Residues: 1-21 <KUZ>
C:Comment: This protein is specific to human P-selectin.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:1-101/Region: V segment
F:16-96/Domain: immunoglobulin homology <IMM>
F:102-113/Region: J segment

Query Match 62.3%; Score 33; DB 2; Length 113;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
||| |||
Db 100 TPRVTGGG 107

RESULT 46
G84604
hypothetical protein At2g21750 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: G84604
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999

```

```

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84604
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <STO>
A:Cross-references: GB:AE002093; NID:g4417270; PIDN:AAD20395.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g21750
A:Map position: 2
C:Superfamily: Arabidopsis thaliana hypothetical protein T22F8.240

Query Match 62.3%; Score 33; DB 2; Length 125;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RVTGGGAM 10
||| |||
Db 37 RVTGGGLM 44

RESULT 47
T36374
probable acetyltransferase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36374
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A:Reference number: Z21573
A:Accession: T36374
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-147 <OLI>
A:Cross-references: EMBL:AL049628; PIDN:CAB40863.1; GSPDB:GN000070; SCOEDB:SCE94.14
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE94.14

Query Match 62.3%; Score 33; DB 2; Length 147;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
||| |||
Db 84 PRVTGGG 90

RESULT 48
S55017
hypothetical protein 2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 16-Feb-1997
C:Accession: S55017
R:Bergetrom, D.E.; Merli, C.A.; Cygan, J.A.; Shelby, R.; Blackman, R.K.
Genetics 139, 1331-1346, 1995
A:Title: Regulatory autonomy and molecular characterization of the Drosophila out at fir
A:Reference number: S55016; MUID:95286060; PMID:7768442
A:Accession: S55017
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-154 <BER>
A:Cross-references: EMBL:L31349
C:Genetics:
A:Gene: FlyBase:oaaf
A:Cross-references: FlyBase:FBgn0011818

Query Match 62.3%; Score 33; DB 2; Length 154;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
||| |||

```

Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
|:|:|:|:  
Db 126 PKVDGGGPM 134

RESULT 39  
RNA-directed RNA polymerase (EC 2.7.7.48) - yeast (Saccharomyces cerevisiae) RNA replicon  
A:Species: Saccharomyces cerevisiae  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jul-2000  
C:Accession: A38149; S27454  
R:Esteban, L.M.; Rodriguez-Cousins, N.; Esteban, R.  
J. Biol. Chem. 267, 10874-10881, 1992  
A:Title: T double-stranded RNA (dsRNA) sequence reveals that T and W dsRNAs form a new R  
A:Reference number: A38149; MUID:92268145; PMID:1587863  
A:Accession: A38149  
A:Molecule type: genomic RNA  
A:Residues: 1-940 <OFF>  
A:Cross-references: EMBL:M86595; NID:g4079660; PIDN:AAC98708.1; PID:g4079661  
A:Note: sequence extracted from NCBI backbone (NCBIN:103889, NCBIP:103891)  
C:Genetics:  
A:Genome: RNA replicon  
C:Keywords: nucleotidyltransferase

Query Match 64.2%; Score 34; DB 2; Length 940;  
Best Local Similarity 66.7%; Pred. No. 3.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
|:|:|:|:  
Db 703 PRVFGGGGL 711

RESULT 40  
T51890  
related to Nup98-Nup96 precursor [imported] - Neurospora crassa  
N:Alternate names: protein B23111.20  
C:Species: Neurospora crassa  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C:Accession: T51890  
R:Schulte, O.; Aign, V.; Hohseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, August 2000  
A:Reference number: Z25858  
A:Accession: T51890  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1844 <SCH>  
A:Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.20  
A:Experimental source: BAC clone B23111; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B23111.20  
A:Map position: 6  
A:Introns: 34/1, 1281/3

Query Match 64.2%; Score 34; DB 2; Length 1844;  
Best Local Similarity 75.0%; Pred. No. 6.8e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
|:|:|:|:  
Db 91 TPAATGGG 98

RESULT 41  
D70575  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: D70575  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: D70575  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-3300 <COL>  
A:Cross-references: GB:Z95324; GB:AL123456; NID:g3261760; PIDN:CAB08587.1; PID:e1299834;  
A:Experimental source: strain H37RV  
A:Gene: PPE

Query Match 64.2%; Score 34; DB 2; Length 3300;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
|:|:|:|:  
Db 173 PTVAGGGAL 181

RESULT 42  
S47208  
T-cell receptor J-alpha wnt1.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 23-Jul-1999  
C:Accession: S47208  
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S40133  
A:Accession: S47208  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-23 <PLA>  
A:Cross-references: EMBL:X71026; NID:G506575; PIDN:CAAS0343.1; PID:g510308  
A:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 62.3%; Score 33; DB 2; Length 23;  
Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
|:|:|:|:  
Db 1 SPGVTGGG 8

RESULT 43  
B37266  
Ig kappa chain V region (2G8) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 21-Jan-2000  
C:Accession: B37266  
R:Ruif-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.  
J. Biol. Chem. 266, 6607-6613, 1991  
A:Title: Heavy and light chain variable region sequences and antibody properties of anti-  
A:Reference number: A38740; MUID:91177923; PMID:1706720  
A:Accession: B37266  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-111 <RUF>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:19-93/Domain: immunoglobulin homology <IMM>

Query Match 62.3%; Score 33; DB 2; Length 111;  
Best Local Similarity 75.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
 C/Accession: F95101  
 R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
 non, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A/Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.  
 A/Reference number: A95000; MUID:21357209; PMID:11463916  
 A/Accession: F95101  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-380 <KUR>  
 A/Cross-references: GB:AE005672; PIDN:AAK75007.1; PID:gi4972353; GSPDB:GN00164; TIGR:SP4  
 A/Experimental source: strain TIGR4  
 C/Genetics:  
 A/Gene: SP0880  
 C/Superfamily: nitrogen fixation protein nifs

Query Match 64.2%; Score 34; DB 2; Length 380;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
 ||:||||  
 Db 221 TPLTGGG 228

RESULT 35  
 G97969  
 hypothetical protein spr0783 [imported] - *Streptococcus pneumoniae* (strain R6)  
 C/Species: *Streptococcus pneumoniae*  
 C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001  
 C/Accession: G97969  
 R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B  
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
 y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
 A/Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
 A/Reference number: A97872; MUID:21429245; PMID:11544234  
 A/Accession: G97969  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-380 <KUR>  
 A/Cross-references: GB:AE007317; PIDN:AAK99587.1; PID:gi5458380; GSPDB:GN00174  
 C/Genetics:  
 A/Gene: spr0783  
 C/Superfamily: nitrogen fixation protein nifs

Query Match 64.2%; Score 34; DB 2; Length 380;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
 ||:||||  
 Db 221 TPLTGGG 228

RESULT 36  
 G70925  
 probable PPE protein - *Mycobacterium tuberculosis* (strain H37Rv)  
 C/Species: *Mycobacterium tuberculosis*  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C/Accession: G70925  
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
 ; Connor, R.; Davies, R.; Devlin, K.; Reltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: G70925

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-408 <COL>  
 A/Cross-references: GB:Z74024; GB:AL123456; NID:g3250700; PIDN:CAA98377.1; PID:el1301025;  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Gene: PPE

Query Match 64.2%; Score 34; DB 2; Length 408;  
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 10  
 ||:||||  
 Db 292 SPSVAGGAV 301

RESULT 37  
 G90481  
 dehydrogenase, probable [imported] - *Sulfolobus solfataricus*  
 C/Species: *Sulfolobus solfataricus*  
 C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C/Accession: G90481  
 R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v  
 jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-NGOC, H.P.; Redder, P.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 Submitted to GenBank, April 2001  
 A/Description: *Sulfolobus solfataricus* complete genome.  
 A/Reference number: A99139  
 A/Accession: G90481  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-409 <KUR>  
 A/Cross-references: GB:AE006641; NID:gi3816397; PIDN:AAK43110.1; GSPDB:GN00155  
 C/Genetics:  
 A/Gene: SSO3008

Query Match 64.2%; Score 34; DB 2; Length 409;  
 Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 10  
 ||:||||  
 Db 180 PRISGGAL 188

RESULT 38  
 G96600  
 protein Fl4J16.24 [imported] - *Arabidopsis thaliana*  
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: G96600  
 R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: G96600  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-440 <STO>  
 A/Cross-references: GB:AE005173; NID:g8778323; PIDN:AAF79332.1; GSPDB:GN00141  
 C/Genetics:  
 A/Gene: Fl4J16.24  
 A/Map position: 1

Query Match 64.2%; Score 34; DB 2; Length 440;

```

A;Molecule type: DNA
A;Residues: 1-288 <BEV>
A;Cross-references: EMBL:AL021749
A;Experimental source: cultivar Columbia; BAC clone F2009
C;Genetics:
A;Map position: 4
A;Introns: 57/3; 96/3; 153/3; 224/3
A;Note: F2009.200
C;Superfamily: erythrocyte band 7 integral membrane protein

Query Match      64.2%; Score 34; DB 2; Length 288;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGAM 10
Db 7 PKIPGGGAI 15

RESULT 30
C75430
uricase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: C75430
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75430
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <WHI>
A;Cross-references: GB:AE001965; GB:AE000513; NID:96458892; PIDN:AAF10733.1; PID:9645889
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1160
A;Map position: 1
C;Superfamily: urate oxidase

Query Match      64.2%; Score 34; DB 2; Length 298;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9
Db 105 PRVTGGGA 112

RESULT 31
AH2629
2'-deoxycytidine 5'-triphosphate deaminase [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH2629
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL41454.1; PID:gi7738778; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: dcd

A;Map position: circular chromosome
Query Match      64.2%; Score 34; DB 2; Length 370;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 195 PNVTTGG 201

RESULT 32
E97411
hypothetical protein AGR_C_764 [imported] - Agrobacterium tumefaciens (strain C58, Cereor
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: E97411
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E97411
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-370 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86246.1; PID:gl5155352; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_764
A;Map position: circular chromosome

Query Match      64.2%; Score 34; DB 2; Length 370;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 195 PNVTTGG 201

RESULT 33
VCVGCN
coat protein - cucumber necrosis virus
C;Species: cucumber necrosis virus
A;Note: host Cucumis sativus (cucumber)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999
C;Accession: JA0131
R;Rochon, D.M.; Tremaine, J.H.
Virology 169, 251-259, 1989
A;Title: Complete nucleotide sequence of the cucumber necrosis virus genome.
A;Reference number: A94391; MUID:89204896; PMID:2705296
A;Accession: JA0131
A;Molecule type: genomic RNA
A;Residues: 1-380 <ROC>
A;Cross-references: GB:M25270; NID:g323338; PIDN:AAA42904.1; PID:g9323341
C;Comment: The genome is a single-stranded, positive-sense RNA. It codes for 33K, 92K, 4(
are core proteins.
C;Superfamily: carnation mottle virus coat protein
C;Keywords: coat protein

Query Match      64.2%; Score 34; DB 1; Length 380;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPVRTGGGAM 10
Db 318 TPSISGGGTI 327

RESULT 34
F95101
aminotransferase, class-V [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae

```

R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, October 1998  
A;Reference number: Z15380  
A;Accession: T04622



```

Db          470  PRVSGG 476

RESULT 20
S57972
hypothetical protein YDR145w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YD8358.02
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999
C;Accession: S57972
R;Murphy, L.; Richards, C.; Harris, D.
submitted to the EMBL Data Library, July 1995
A;Reference number: S57971
A;Accession: S57972
A;Molecule type: DNA
A;Residues: 1-539 <MUR>
A;Cross-references: EMBL:Z50046; NID:g989393; PIDN:CAA90368.1; PID:g989395; MIPS:YDR145w
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:TAF61
A;Cross-references: SGD:S0002552; MIPS:YDR145w
A;Map position: 4r

Query Match          66.0%; Score 35; DB 2; Length 539;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy          2  PRVTGGGAM 10
| : | | | |
Db          383  PTITGGSAM 391

RESULT 21
E70820
hypothetical glycine-rich protein Rv0977 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: E70820
R;Cole, S.T.; Broach, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70820
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-923 <COL>
A;Cross-references: GB:AL021999; GB:AL123456; NID:g3261538; PIDN:CAA17576.1; PID:g291693
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0977
C;Superfamily: elastin

Query Match          66.0%; Score 35; DB 2; Length 923;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy          1  TPRVTGGAM 10
| : | | | |
Db          887  TPRVTSGGVM 896

RESULT 22
T36410
probable polyketide synthetase SCF34.11c [similarity] - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Nov-2000
C;Accession: T36410
R;Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21606

A;Accession: T36410
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3643 <SAU>
A;Cross-references: EMBL:AL109974; PIDN:CAB53322.1; GSPDB:GN00070; SCOEDB:SCF34.11c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCF34.11c
C;Superfamily: alpha-aminoacyl-cysteine-valine synthetase; acetate-CoA ligase homology
N;Alternate names: carrier protein; phosphopantetheine; phosphoprotein
C;Keywords: carrier protein; acetate-CoA ligase homology <ACL1>
F;183-630/Domain: acetate-CoA ligase homology <ACP1>
F;645-712/Domain: acyl carrier protein homology <ACP1>
F;1671-2121/Domain: acetate-CoA ligase homology <ACL2>
F;2140-2207/Domain: acyl carrier protein homology <ACP2>
F;3101-3550/Domain: acetate-CoA ligase homology <ACL3>
F;3569-3639/Domain: acyl carrier protein homology <ACP3>
F;677,2172,3603/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match          66.0%; Score 35; DB 2; Length 3643;
Best Local Similarity 85.7%; Pred. No. 8.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          1  TPRVTGG 7
| : | | | |
Db          3163  TPRLTGG 3169

RESULT 23
T49671
hypothetical protein B8B20.250 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C;Accession: T49671
R;Schulte, U.; Align, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-211 <SCH>
A;Cross-references: EMBL:AL355933; GSPDB:GN00116; NCSP:B8B20.250
A;Experimental source: BAC clone B8B20; strain OR74A
C;Genetics:
A;Gene: NCSP:B8B20.250
A;Map position: 6
C;Superfamily: Neurospora crassa hypothetical protein B8B20.250

Query Match          64.2%; Score 34; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy          2  PRVTGGGAM 10
| : | | | |
Db          55  PRVYGGGAV 53

RESULT 24
S64065
hypothetical protein YGL061c - Yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G3432
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C;Accession: S64065
R;Feuermann, M.; Potier, S.; Souciet, J.L.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64044
A;Accession: S64065
A;Molecule type: DNA
A;Residues: 1-247 <FEU>
A;Cross-references: EMBL:Z72583; NID:gl322563; PID:g243286; GSPDB:GN00007;
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:DUO1; MIPS:YGL061c

```

T18583  
 glucosylceramidase (EC 3.2.1.45) precursor - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T18583; T20789  
 R:Smyle, R.  
 submitted to the EMBL Data Library, August 1998  
 A:Reference number: Z18999  
 A:Accession: T18583  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-522 <W12>  
 A:Cross-references: EMBL:AL031254; PIDN:CAA02083.1; GSPDB:GN00022; CESP:FL11E6.1  
 A:Experimental source: clone 4R79  
 R:Matthews, L.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19324  
 A:Accession: T20789  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-522 <W12>  
 A:Cross-references: EMBL:Z81058; PIDN:CAB02924.1; GSPDB:GN00022; CESP:FL11E6.1  
 A:Experimental source: clone File6  
 C:Genetics:  
 A:Gene: CESP:File6.1  
 A:Map position: 4  
 A:Introns: 19/1; 89/1; 183/3; 239/2; 364/3; 394/1; 447/2  
 C:Superfamily: glucosylceramidase  
 C:Keywords: glycoprotein; glycosidase; hydrolase; lysosome; membrane bound; sphingolipid  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-522/Product: glucosylceramidase #status predicted <MPT>  
 F:42,86,90,274,428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.0%; Score 35; DB 2; Length 522;  
 Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 ||:|||||  
 DB 214 TGRMRGGGAM 223

RESULT 19  
 H86228  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: H86228  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: H86228  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-524 <STO>  
 A:Cross-references: GB:AE005172; NID:G3482928; PIDN:AAC33213.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 Query Match 66.0%; Score 35; DB 2; Length 524;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 ||:|||||

RESULT 16  
 A53138  
 gas1 homolog - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000  
 C:Accession: A53138  
 R:Del Sal, G.; Collavin, L.; Ruaro, M.E.; Edomi, P.; Saccone, S.; della Valle, G.; Schne  
 Proc. Natl. Acad. Sci. U.S.A. 91, 1948-1952, 1994  
 A:Title: Structure, function, and chromosome mapping of the growth-suppressing human hom  
 A:Reference number: A53138; MUID:94173926; PMID:8127893  
 A:Accession: A53138  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-345 <RES>  
 A:Cross-references: GB:L13698; NID:G472859; PIDN:AAA72368.1; PID:G472860

Query Match 66.0%; Score 35; DB 2; Length 345;  
 Best Local Similarity 75.0%; Pred. No. 88;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 ||:|||||  
 DB 148 PRVTGGGA 155

RESULT 17  
 JC7280  
 cytokine receptor-like molecule-2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C:Accession: JC7280  
 R:Hiroyama, T.; Iwama, A.; Morita, Y.; Nakamura, Y.; Shibuya, A.; Nakauchi, H.  
 Biochem. Biophys. Res. Commun. 272, 224-229, 2000  
 A:Title: Molecular cloning and characterization of CRLM-2, a novel type I cytokine recep  
 A:Reference number: JC7280  
 A:Contents: Embryo  
 A:Accession: JC7280  
 A:Molecule type: mRNA  
 A:Residues: 1-359 <CHIR>  
 A:Cross-references: DDBJ:AB039945  
 C:Genetics:  
 A:Gene: Crim-2  
 C:Keywords: cytokine; embryo; receptor; signal transduction; transmembrane protein

Query Match 66.0%; Score 35; DB 2; Length 359;  
 Best Local Similarity 77.8%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 PRVTGGGAM 10  
 ||:|||||  
 DB 333 PRVTGGGAM 341

RESULT 18

T18583  
 glucosylceramidase (EC 3.2.1.45) precursor - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T18583; T20789  
 R:Smyle, R.  
 submitted to the EMBL Data Library, August 1998  
 A:Reference number: Z18999  
 A:Accession: T18583  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-522 <W12>  
 A:Cross-references: EMBL:AL031254; PIDN:CAA02083.1; GSPDB:GN00022; CESP:FL11E6.1  
 A:Experimental source: clone 4R79  
 R:Matthews, L.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19324  
 A:Accession: T20789  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-522 <W12>  
 A:Cross-references: EMBL:Z81058; PIDN:CAB02924.1; GSPDB:GN00022; CESP:FL11E6.1  
 A:Experimental source: clone File6  
 C:Genetics:  
 A:Gene: CESP:File6.1  
 A:Map position: 4  
 A:Introns: 19/1; 89/1; 183/3; 239/2; 364/3; 394/1; 447/2  
 C:Superfamily: glucosylceramidase  
 C:Keywords: glycoprotein; glycosidase; hydrolase; lysosome; membrane bound; sphingolipid  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-522/Product: glucosylceramidase #status predicted <MPT>  
 F:42,86,90,274,428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.0%; Score 35; DB 2; Length 522;  
 Best Local Similarity 70.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 ||:|||||  
 DB 214 TGRMRGGGAM 223

RESULT 19  
 H86228  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: H86228  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: H86228  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-524 <STO>  
 A:Cross-references: GB:AE005172; NID:G3482928; PIDN:AAC33213.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 Query Match 66.0%; Score 35; DB 2; Length 524;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 ||:|||||

RESULT 16  
 A53138  
 gas1 homolog - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000  
 C:Accession: A53138  
 R:Del Sal, G.; Collavin, L.; Ruaro, M.E.; Edomi, P.; Saccone, S.; della Valle, G.; Schne  
 Proc. Natl. Acad. Sci. U.S.A. 91, 1948-1952, 1994  
 A:Title: Structure, function, and chromosome mapping of the growth-suppressing human hom  
 A:Reference number: A53138; MUID:94173926; PMID:8127893  
 A:Accession: A53138  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-345 <RES>  
 A:Cross-references: GB:L13698; NID:G472859; PIDN:AAA72368.1; PID:G472860

Query Match 66.0%; Score 35; DB 2; Length 345;  
 Best Local Similarity 75.0%; Pred. No. 88;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 ||:|||||  
 DB 148 PRVTGGGA 155

RESULT 17  
 JC7280  
 cytokine receptor-like molecule-2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C:Accession: JC7280  
 R:Hiroyama, T.; Iwama, A.; Morita, Y.; Nakamura, Y.; Shibuya, A.; Nakauchi, H.  
 Biochem. Biophys. Res. Commun. 272, 224-229, 2000  
 A:Title: Molecular cloning and characterization of CRLM-2, a novel type I cytokine recep  
 A:Reference number: JC7280  
 A:Contents: Embryo  
 A:Accession: JC7280  
 A:Molecule type: mRNA  
 A:Residues: 1-359 <CHIR>  
 A:Cross-references: DDBJ:AB039945  
 C:Genetics:  
 A:Gene: Crim-2  
 C:Keywords: cytokine; embryo; receptor; signal transduction; transmembrane protein

Query Match 66.0%; Score 35; DB 2; Length 359;  
 Best Local Similarity 77.8%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 PRVTGGGAM 10  
 ||:|||||  
 DB 333 PRVTGGGAM 341

RESULT 18

QY 3 RVTGGGAM 10  
 Db 134 RITGGGAI 141

## RESULT 11

T01663  
 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) IIb - maize  
 N;Alternate names: starch branching enzyme IIb  
 C;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 18-Jun-1999  
 C;Species: Zea mays (maize)  
 C;Accession: T01663  
 R;Kim, K.N.; Fisher, D.K.; Gao, M.; Guiltinan, M.J.  
 submitted to the EMBL Data Library, June 1998  
 A;Description: Molecular cloning and characterization of the amylose-extender gene encoding  
 A;Reference number: Z14387  
 A;Accession: T01663  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-799 <KIM>  
 A;Cross-references: EMBL:AF072725; NID:g3511235; PIDN:AAC33764.1; PID:g3511236  
 A;Experimental source: strain B73  
 C;Genetics:  
 A;Gene: ae  
 A;Introns: 38/1; 86/3; 138/2; 171/2; 185/3; 205/3; 232/3; 271/3; 299/3; 340/2; 380/2; 42

C;Function:  
 A;Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-D-g  
 A;Pathway: glycogen/starch biosynthesis  
 C;Superfamily: 1,4-alpha-glucan branching enzyme  
 C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 67.9%; Score 36; DB 2; Length 799;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
 Db 17 PRLTGGG 23

## RESULT 12

T02981  
 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) II - maize  
 N;Alternate names: starch branching enzyme II  
 C;Species: Zea mays (maize)  
 C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 18-Jun-1999  
 C;Accession: T02981  
 R;Fisher, D.K.; Boyer, C.D.; Hannah, L.C.  
 Plant Physiol. 102, 1045-1046, 1993  
 A;Title: Starch branching enzyme II from maize endosperm.  
 A;Reference number: Z14808; MUID:94105320; PMID:8278524  
 A;Accession: T02981  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-799 <FIS>  
 A;Cross-references: EMBL:D08065; NID:g168482; PIDN:AAA18571.1; PID:g168483  
 A;Experimental source: cultivar W64Ax182E  
 C;Function:  
 A;Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-D-g

Query Match 67.9%; Score 36; DB 2; Length 799;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
 Db 17 PRLTGGG 23

## RESULT 13

AH2893  
 conserved hypothetical protein Atu2585 [imported] - Agrobacterium tumefaciens (strain C58  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C;Accession: AH2893  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I  
 ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AH2893  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-223 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL43566.1; PID:g17741080; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atu2585  
 A;Map position: circular chromosome

Query Match 66.0%; Score 35; DB 2; Length 223;  
 Best Local Similarity 70.0%; Pred. No. 58;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
 Db 119 TPRDTGGGSM 128

## RESULT 14

C97669  
 hypothetical protein AGR\_C\_4682 [imported] - Agrobacterium tumefaciens (strain C58, Cerc  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C;Accession: C97669  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: C97669  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-304 <KUR>  
 A;Cross-references: GB:AE007869; PIDN:AAK88308.1; PID:g15157780; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C\_4682  
 A;Map position: circular chromosome

Query Match 66.0%; Score 35; DB 2; Length 304;  
 Best Local Similarity 70.0%; Pred. No. 78;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
 Db 200 TPRDTGGGSM 209

## RESULT 15

D72243  
 rod shape-determining protein MreB - Thermotoga maritima (strain MSBB)  
 C;Species: Thermotoga maritima  
 C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C;Accession: D72243  
 R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 C.M.

Nature 399, 323-329, 1999  
 A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A;Reference number: A72200; MUID:99287316; PMID:10360571

S25281  
 gltF protein precursor - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 01-Mar-2002  
 C:Accession: S25281; H65112  
 R:Castano, I.; Flores, N.; Valle, F.; Covarrubias, A.A.; Bolivar, F.  
 Mol Microbiol. 6, 2733-2741, 1992  
 A:Title: gltF, a member of the gltBDF operon of Escherichia coli, is involved in nitroge  
 A:Reference number: S25281; MUID:93078627; PMID:1447980  
 A:Accession: S25281  
 A:Molecule type: DNA  
 A:Residues: 1-254 <CAS>  
 A:Cross-references: EMBL:M74162; NID:gl46213; PIDN:AAA23909.1; PID:gl46214  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: H65112  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-254 <BLAT>  
 A:Cross-references: GB:AE000401; GB:U00096; NID:gl789607; PIDN:AAC76246.1; PID:gl789608;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: gltF  
 C:Keywords: transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:27-254/Product: gltF protein #status predicted <MAT>  
 F:213-229/Domain: transmembrane #status predicted <TM>  
 Query Match 67.9%; Score 36; DB 2; Length 254;  
 Best Local Similarity 87.5%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TPRVTGGG 8  
 ||| |||||  
 Db 45 TPVTGGG 52  
 RESULT 7  
 T46683  
 probable lipocate protein ligase snop [imported] - Streptomyces nogalater  
 C:Species: Streptomyces nogalater  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 18-Feb-2000  
 C:Accession: T46683  
 R:Ylilhonko, K.P.J.  
 submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z23126  
 A:Accession: T46683  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-354 <YLI>  
 A:Cross-references: EMBL:AJ224512; PIDN:CAB59005.1  
 A:Experimental source: ATCC 27451  
 C:Genetics:  
 A:Gene: snop  
 Query Match 67.9%; Score 36; DB 2; Length 354;  
 Best Local Similarity 75.0%; Pred. No. 60;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RVTGGGAM 10  
 ||| |||||  
 Db 171 RISGGGAM 178  
 RESULT 8  
 T28693  
 hypothetical protein - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T28693

R;Parkhill, J.; Bentley, S.D.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, January 1999  
 A:Reference number: Z20512  
 A:Accession: T28693  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-388 <PAR>  
 A:Cross-references: EMBL:AL023496; NID:el292348; PID:el292360; PIDN:CAA18910.1

Query Match 67.9%; Score 36; DB 2; Length 388;  
 Best Local Similarity 75.0%; Pred. No. 65;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RVTGGGAM 10  
 ||| |||||  
 Db 205 RISGGGAM 212

## RESULT 9

T02294  
 hypothetical protein T13D8.31 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 10-Dec-1999  
 C:Accession: T02294  
 R;Vyotsakia, V.S.; Schwartz, J.R.; Kwan, A.; Toriumi, M.; Yu, G.; Oji, O.; Liu, S.; Li,  
 rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.  
 submitted to the EMBL Data Library, June 1998  
 A:Description: Arabidopsis thaliana chromosome 1 BAC T13D8 sequence.  
 A:Reference number: Z14649  
 A:Accession: T02294  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-404 <VYS>  
 A:Cross-references: EMBL:AC004473; NID:g3108025; PID:g3249085  
 C:Genetics:  
 A:Map position: 1  
 A:Introns: 15/1; 70/1; 90/2; 145/3; 185/3; 217/2; 248/3; 292/3; 330/2  
 A:Note: T13D8.31  
 C:Superfamily: Saccharomyces hypothetical protein YDR531w  
 Query Match 67.9%; Score 36; DB 2; Length 404;  
 Best Local Similarity 75.0%; Pred. No. 68;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 ||| |||||  
 Db 105 PRATGGGA 112

## RESULT 10

C70441  
 lipocate-protein ligase A - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
 C:Accession: C70441  
 R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove  
 V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666; PMID:9537320  
 A:Accession: C70441  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-788 <AQF>  
 A:Cross-references: GB:AE000748; NID:g2983960; PIDN:AAC07512.1; PID:g2983965; GB:AE00065;  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: lplA  
 Query Match 67.9%; Score 36; DB 2; Length 788;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;



395 30 56.6 583 2 C37855  
 396 30 56.6 592 2 B81009  
 397 30 56.6 592 2 H82031  
 398 30 56.6 610 2 S59394  
 399 30 56.6 613 2 A82834  
 400 30 56.6 618 2 A33588  
 401 30 56.6 619 2 T48557  
 402 30 56.6 625 2 T02847  
 403 30 56.6 627 2 T02415  
 404 30 56.6 631 2 E69115  
 405 30 56.6 646 2 H1378  
 406 30 56.6 653 2 A60961  
 407 30 56.6 657 2 T48228  
 408 30 56.6 680 2 H87091  
 409 30 56.6 692 2 C55926  
 410 30 56.6 706 2 F87683  
 411 30 56.6 738 2 D70680  
 412 30 56.6 766 2 G69550  
 413 30 56.6 776 2 A55448  
 414 30 56.6 789 2 I39550  
 415 30 56.6 791 2 A56241  
 416 30 56.6 795 2 H70934  
 417 30 56.6 798 2 T33022  
 418 30 56.6 802 2 T23295  
 419 30 56.6 860 2 T23296  
 420 30 56.6 863 2 A53034  
 421 30 56.6 891 1 JN0867  
 422 30 56.6 894 2 PNO667  
 423 30 56.6 905 2 E64058  
 424 30 56.6 943 2 T35497  
 425 30 56.6 947 2 T23107  
 426 30 56.6 1051 2 AG3455  
 427 30 56.6 1052 2 C64221  
 428 30 56.6 1099 2 A59300  
 429 30 56.6 1323 2 T00037  
 430 30 56.6 1366 2 T35985  
 431 30 56.6 1503 2 T18266  
 432 30 56.6 1524 2 T30337  
 433 30 56.6 1551 2 A83488  
 434 30 56.6 1662 2 T18540  
 435 30 56.6 2116 2 C86926  
 436 30 56.6 2118 2 S72705  
 437 30 56.6 2150 2 T32497  
 438 30 56.6 2793 2 B90784  
 439 30 56.6 2806 2 D85644  
 440 30 56.6 3716 2 E70969  
 441 30 56.6 6831 2 A88852  
 442 30 56.6 6839 2 S57242  
 443 30 56.6 7160 2 T27935  
 444 30 55.7 303 2 A84344  
 445 30 55.7 320 2 T21452  
 446 30 55.7 334 2 B72393  
 447 29.5 55.7 34 2 S47211  
 448 29.5 55.7 25 2 S07250  
 449 29.5 54.7 25 2 T46085  
 450 29.5 54.7 49 2 C81855  
 451 29.5 54.7 55 2 S29941  
 452 29.5 54.7 90 2 AF2322  
 453 29.5 54.7 97 2 C87685  
 454 29.5 54.7 104 2 A48677  
 455 29.5 54.7 107 2 B48677  
 456 29.5 54.7 107 2 B48677  
 457 29.5 54.7 107 2 D48677  
 458 29.5 54.7 111 2 AH2530  
 459 29.5 54.7 126 2 AG1125  
 460 29.5 54.7 133 2 E75215  
 461 29.5 54.7 134 1 WKBPT5  
 462 29.5 54.7 136 2 E30593  
 463 29.5 54.7 138 2 B46348  
 464 29.5 54.7 140 2 F86288  
 465 29.5 54.7 142 2 T52145  
 466 29.5 54.7 145 2 H87436  
 467 29.5 54.7 146 2 S53010

phosphogluconate d  
 BirA protein/Bvg a  
 probable biotin-[a  
 protein kinase RCK  
 hypothetical prote  
 mismatch repair pr  
 hypothetical prote  
 probable membrane  
 probable homeodoma  
 phage infection pr  
 conserved hypothet  
 leishmanolysin (EC  
 probable protein k  
 probable NAD synth  
 DNA binding protei  
 peptidase M13 fami  
 ATP-dependent RNA  
 Ah receptor nuclea  
 aryl hydrocarbon r  
 aryl hydrocarbon r  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 gag polyprotein -  
 peroxinectin-like  
 peroxinectin-like  
 hemopexin-heme com  
 hypothetical prote  
 hypothetical prote  
 multidrug resistan  
 hypothetical l14K  
 myosin-II - mouse  
 hypothetical prote  
 probable large Pro  
 cycloimmun-oligosa  
 polyporein - Afri  
 gramicidin S biosy  
 mofA protein precu  
 probable myoceros  
 myocerosate synth  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 probable PPE prote  
 protein unc-22 [im  
 twitchin [similar  
 hypothetical prote  
 NADH dehydrogenase  
 oligopeptide ABC t  
 T-cell receptor J-  
 hypothetical prote  
 hypothetical prote  
 hypothetical prote  
 late embryogenesis  
 hypothetical prote  
 hypothetical prote  
 Ig kappa chain V-J  
 Ig kappa chain V-J  
 Ig kappa chain V-J  
 Ig kappa chain V-J  
 hypothetical prote  
 B. subtilis yxh p  
 glyoxalase I relat  
 gene A2 protein -  
 outer membrane pro  
 gene A2-A3 protein  
 F9L1.35 protein -  
 ribosomal protein  
 hypothetical prote  
 RCc2 protein - ric

468 29 54.7 146 2 S53011  
 469 29 54.7 151 2 S21600  
 470 29 54.7 155 1 TIZM1  
 471 29 54.7 159 2 H84224  
 472 29 54.7 163 2 AH3002  
 473 29 54.7 163 2 T47559  
 474 29 54.7 166 2 H98280  
 475 29 54.7 168 2 H64237  
 476 29 54.7 175 2 AB3345  
 477 29 54.7 176 2 AG2817  
 478 29 54.7 178 2 H97595  
 479 29 54.7 180 2 C81235  
 480 29 54.7 182 2 S22412  
 481 29 54.7 185 2 C87646  
 482 29 54.7 195 2 AC0699  
 483 29 54.7 206 2 G86279  
 484 29 54.7 208 2 E84316  
 485 29 54.7 209 2 T22589  
 486 29 54.7 210 2 JH0794  
 487 29 54.7 212 2 AE1838  
 488 29 54.7 215 2 G49343  
 489 29 54.7 226 2 AF2965  
 490 29 54.7 229 2 S72165  
 491 29 54.7 232 2 D70537  
 492 29 54.7 234 2 C83587  
 493 29 54.7 235 2 T36923  
 494 29 54.7 236 2 B91228  
 495 29 54.7 236 2 A86075  
 496 29 54.7 236 2 S40817  
 497 29 54.7 241 2 D87135  
 498 29 54.7 247 2 B90500  
 499 29 54.7 248 2 S77172  
 500 29 54.7 251 2 T34168

ALIGNMENTS

RESULT 1

WMBE65  
 6SK lower matrix phosphoprotein - human cytomegalovirus (strain AD169)  
 N:Alternate names: protein UL83  
 C:Species: human cytomegalovirus, human herpesvirus 5  
 C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 16-Jul-1999  
 C:Accession: A26793; S09847  
 R:Rueger, B.; Klages, S.; Walla, B.; Albrecht, J.; Fleckenstein, B.; Tomlinson, P.; Barri  
 J. Virol. 61, 446-453, 1987  
 A:Title: Primary structure and transcription of the genes coding for the two virion phosph  
 A:Reference number: A93028; MUID:87112940; PMID:3027374  
 A:Accession: A26793  
 A:Molecule type: DNA  
 A:Residues: 1-561 <RUE>  
 A:Cross-references: GB:M15120  
 R:Cher, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; I  
 M.; Barrell, B.G.  
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
 A:Reference number: S09749; MUID:90269039; PMID:2161319  
 A:Accession: S09847  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-561 <CHE>  
 A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35357.1; PID:g1780861  
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
 C:Genetics: 93/3  
 A:Introns: 93/3  
 C:Superfamily: cytomegalovirus lower matrix phosphoprotein  
 C:Keywords: matrix protein; phosphoprotein  
 Query Match 100.0%; Score 53; DB 1; Length 561;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

249	31	58.5	908	2	AE2675	pyruvate, orthophos	322	30	56.6	322	2	AF0709	succinylglutamate
250	31	58.5	915	2	AC3477	DNA mismatch repai	323	30	56.6	329	2	AD1170	lipoate-protein li
251	31	58.5	918	2	T02759	hypothetical prote	324	30	56.6	329	2	AF1527	lipoate-protein li
252	31	58.5	933	2	C97457	pyruvate, phosphat	325	30	56.6	329	2	H64705	conserved hypothet
253	31	58.5	1009	1	DJBEM2	DNA-directed DNA p	326	30	56.6	329	2	H71815	hypothetical prote
254	31	58.5	1054	2	UC7294	alpha integrin -	327	30	56.6	330	2	C83735	lipoate-protein li
255	31	58.5	1189	2	SE6852	hypothetical prote	328	30	56.6	331	1	G69830	lipoate-protein li
256	31	58.5	1208	2	T23467	hypothetical prote	329	30	56.6	331	2	AC1191	lipoate-protein li
257	31	58.5	1279	2	T17194	protoporphyrin IX	330	30	56.6	331	2	AC1549	lipoate-protein li
258	31	58.5	1374	2	D72593	hypothetical prote	331	30	56.6	333	2	PN0441	mreB protein - Bac
259	31	58.5	1402	2	I46707	translation initia	332	30	56.6	334	2	C84117	cell shape determi
260	31	58.5	1733	1	BA5344	probable nuclear a	333	30	56.6	334	2	A82098	thiamin-monophosph
261	31	58.5	1958	2	B40505	hypothetical prote	334	30	56.6	335	2	C69660	cell-shape determi
262	31	58.5	2044	2	T13704	still life protein	335	30	56.6	337	2	D69660	cell-shape determi
263	31	58.5	2061	2	T13751	transcription fact	336	30	56.6	337	2	AD1268	cell-shape determi
264	31	58.5	2204	2	A70524	probable PPE prote	337	30	56.6	337	2	AF1630	hypothetical prote
265	31	58.5	2761	2	T29285	hypothetical prote	338	30	56.6	340	2	A89798	stage V sporulatio
266	31	58.5	3519	2	S43048	polyketide synthas	339	30	56.6	342	2	A97184	lipoate-protein li
267	30.5	57.5	500	2	S55785	nucleolar protein	340	30	56.6	344	2	C90567	rod shape-determin
268	30.5	57.5	632	2	AE2382	acetylhydroxy acid	341	30	56.6	345	2	F83084	rod shape-determin
269	30	56.6	17	2	S57514	T cell receptor be	342	30	56.6	346	2	G84028	cell-shape determi
270	30	56.6	55	2	PQ0829	E2/NS1 protein (CD	343	30	56.6	346	2	C81446	homolog of E. coli
271	30	56.6	111	2	T14306	glycine-rich prote	344	30	56.6	347	2	AG2840	rod shape-determin
272	30	56.6	112	2	T09060	hypothetical prote	345	30	56.6	347	2	E71826	rod shape-determin
273	30	56.6	121	2	T48820	hypothetical prote	346	30	56.6	347	2	E64591	rod shape-determin
274	30	56.6	135	2	G72701	hypothetical prote	347	30	56.6	347	2	F87440	rod shape-determin
275	30	56.6	136	2	A71452	glyoxalase I relat	348	30	56.6	347	2	AC0446	rod shape-determin
276	30	56.6	137	2	B41047	exsB protein - Pse	349	30	56.6	347	2	AC0912	rod shape-determin
277	30	56.6	137	2	E83433	exoenzyme S synth	350	30	56.6	347	2	AC0912	rod shape-determin
278	30	56.6	137	2	S24074	envelope protein -	351	30	56.6	351	2	F83419	probable permease
279	30	56.6	138	2	JC7676	cystatin 1 - wheat	352	30	56.6	360	2	F83419	rod shape-determin
280	30	56.6	142	2	JC7676	probable membrane	353	30	56.6	361	2	BVECEB	rod shape-determin
281	30	56.6	155	2	T35626	hypothetical prote	354	30	56.6	367	1	BVECEB	regulator of ftsI
282	30	56.6	155	2	G72580	hypothetical prote	355	30	56.6	367	2	C91144	hypothetical prote
283	30	56.6	163	2	T36985	hypothetical prote	356	30	56.6	367	2	F85989	hypothetical prote
284	30	56.6	172	2	T40835	aspartokinase beta	357	30	56.6	368	2	T15492	hypothetical prote
285	30	56.6	180	2	T06718	hypothetical prote	358	30	56.6	369	2	F82896	hypothetical prote
286	30	56.6	180	2	A42371	hypothetical prote	359	30	56.6	370	2	E71221	hypothetical prote
287	30	56.6	190	2	PQ0019	adenosinetriphosph	360	30	56.6	378	2	E64044	rod shape-determin
288	30	56.6	200	2	T44143	DR3 protein (impor	361	30	56.6	378	2	T07007	probable 1-phospha
289	30	56.6	200	2	T43953	hypothetical prote	362	30	56.6	384	2	S51796	vasodilator-stimul
290	30	56.6	201	2	S55000	alpha-chitin bindi	363	30	56.6	390	2	151419	transcription fact
291	30	56.6	209	2	B87627	hypothetical prote	364	30	56.6	401	2	A57690	aerolysin precurs
292	30	56.6	211	2	G37376	hypothetical prote	365	30	56.6	413	2	D70137	cell division prot
293	30	56.6	225	2	S36621	probable drought-i	366	30	56.6	425	2	H64185	cell division cofact
294	30	56.6	230	2	T45365	ribosomal protein	367	30	56.6	425	2	H64185	iron-sulfur cofact
295	30	56.6	232	2	A54501	sporulated oocyt	368	30	56.6	427	2	G71272	lipopolysaccharide
296	30	56.6	232	2	A10219	flagellar basal-bo	369	30	56.6	434	2	H72358	transcription fact
297	30	56.6	239	2	F84189	hypothetical prote	370	30	56.6	448	2	T37053	hypothetical prote
298	30	56.6	250	2	S22004	gamma-Tip protein	371	30	56.6	449	1	S30205	transcription fact
299	30	56.6	251	2	C70521	1-acylglycerol-3-p	372	30	56.6	451	1	A40168	transcription fact
300	30	56.6	261	2	B70521	1-acylglycerol-3-p	373	30	56.6	451	1	A40168	hypothetical prote
301	30	56.6	264	2	T03974	anthocyanin biosyn	374	30	56.6	453	2	F72499	probable Vi polysa
302	30	56.6	264	2	T03974	anthocyanin biosyn	375	30	56.6	455	2	F71153	probable 3-oxoacid
303	30	56.6	269	2	S74919	hypothetical prote	376	30	56.6	464	2	H75538	probable 3-oxoacid
304	30	56.6	271	2	T03715	anthocyanin biosyn	377	30	56.6	498	1	S47801	L-xylulokinase (EC
305	30	56.6	272	2	T35473	hypothetical prote	378	30	56.6	498	2	G81081	ferredoxin, 4Fe-4S
306	30	56.6	272	2	H84178	lipoate-protein li	379	30	56.6	498	2	H81861	probable ferredoxi
307	30	56.6	281	2	C55519	hypothetical prote	380	30	56.6	498	2	JC4727	mob protein A - Er
308	30	56.6	283	2	H42600	probable octopine-	381	30	56.6	503	2	S04790	mobilization prote
309	30	56.6	283	2	A13230	hypothetical prote	382	30	56.6	509	2	JQ0535	oligo-1,6-glucosid
310	30	56.6	286	2	BA5632	merozoite surface	383	30	56.6	515	1	NNBS1	anchranilate synth
311	30	56.6	288	2	D75286	serine proteinase	384	30	56.6	519	2	S23796	modifier-3 protein
312	30	56.6	290	2	G72203	sugar ABC transpor	385	30	56.6	523	2	B83629	probable ATP-bindi
313	30	56.6	297	2	H44020	hypothetical prote	386	30	56.6	530	2	AF0051	probable carbohydr
314	30	56.6	298	2	A89009	protein T27C4.1 (i	387	30	56.6	530	2	AF0051	probable sugar kin
315	30	56.6	300	2	S47764	dipeptide transpor	388	30	56.6	530	2	AF0051	probable sugar kin
316	30	56.6	300	2	F91181	dipeptide transpor	389	30	56.6	532	2	C95916	capsid protein - N
317	30	56.6	300	2	B86028	dipeptide transpor	390	30	56.6	533	2	S36477	probable sugar kin
318	30	56.6	300	2	A10486	dipeptide transpor	391	30	56.6	534	2	T01500	cytokinin oxidase
319	30	56.6	300	2	AG0983	dipeptide transpor	392	30	56.6	534	2	T51929	cytokinin oxidase
320	30	56.6	301	2	G87563	conserved hypothet	393	30	56.6	551	1	NRECE3	colicin E3 (EC 3.1
321	30	56.6	306	2	T08504	trbG protein - Ent	394	30	56.6	557	2	C72686	probable thermosom



103	32	60.4	264	2	E90500	lipoate-protein li	176	31	58.5	274	2	H83525	conserved hypothet
104	32	60.4	264	2	B72462	probable lipoate-p	177	31	58.5	275	1	E36516	cis-1,2-dihydroben
105	32	60.4	271	2	S23239	chymotrypsin (EC 3	178	31	58.5	289	2	S74632	hypothetical prote
106	32	60.4	309	2	AD3594	transporter, dme f	179	31	58.5	296	2	C82605	conjugal transfer
107	32	60.4	317	2	AS1572	moca protein - Rhi	180	31	58.5	296	2	B82866	conjugal transfer
108	32	60.4	328	2	C89871	hypothetical prote	181	31	58.5	298	2	UC1448	collagen col-34 -
109	32	60.4	329	2	D95134	lipoate-protein li	182	31	58.5	299	2	T29956	hypothetical prote
110	32	60.4	329	2	G98002	lipoate protein li	183	31	58.5	311	2	T07838	chitinase (EC 3.2.
111	32	60.4	333	2	I40531	cell shape determi	184	31	58.5	311	2	C72594	hypothetical prote
112	32	60.4	336	1	Q0BE7	BFRF1 protein - hu	185	31	58.5	311	2	F72502	hypothetical prote
113	32	60.4	345	2	H90515	lipoate-protein li	186	31	58.5	325	1	Q0BE84	UL76 protein - num
114	32	60.4	358	2	S41640	ribosomal protein	187	31	58.5	338	2	R82204	lipoate-protein li
115	32	60.4	362	2	S45887	ribosomal protein	188	31	58.5	346	2	T49148	nitrilase (EC 3.5.
116	32	60.4	362	2	S50993	ribosomal protein	189	31	58.5	364	2	AE2994	glycosyltransferas
117	32	60.4	363	2	T40797	60s ribosomal prot	190	31	58.5	368	2	E83559	hypothetical prote
118	32	60.4	363	2	D69502	rod shape-determi	191	31	58.5	369	2	B72327	dnaj protein - The
119	32	60.4	378	2	S76183	hypothetical prote	192	31	58.5	369	2	D81984	riboflavin bifunct
120	32	60.4	379	2	AF2409	mannosyl transfera	193	31	58.5	369	2	H87633	threonine aldolase
121	32	60.4	410	2	E75080	3-phosphoglycerate	194	31	58.5	370	2	H70583	phosphate-binding
122	32	60.4	410	2	D71065	probable phosphogl	195	31	58.5	370	2	F83559	hypothetical prote
123	32	60.4	410	2	S68188	phosphoglycerate k	196	31	58.5	371	2	F75301	conserved hypothet
124	32	60.4	417	2	H82215	serine transporter	197	31	58.5	372	2	C97791	nifs protein homol
125	32	60.4	427	2	A86493	polymorphic outer	198	31	58.5	372	2	E98289	mannosyltransferas
126	32	60.4	432	2	S26697	aprE protein - Pse	199	31	58.5	380	1	GNVSMB	genome polyprotein
127	32	60.4	432	2	G83489	alkaline proteinas	200	31	58.5	380	2	AE0377	probable mannosyl
128	32	60.4	434	2	T37521	nifs homolog - fis	201	31	58.5	394	2	T26968	hypothetical prote
129	32	60.4	441	2	E72579	hypothetical prote	202	31	58.5	404	2	T20055	hypothetical prote
130	32	60.4	444	2	AC3127	two component sens	203	31	58.5	410	2	A71652	iron-sulfur cofact
131	32	60.4	444	2	F98160	probable two-compo	204	31	58.5	410	2	B97791	nifs protein homol
132	32	60.4	457	2	T29116	hypothetical prote	205	31	58.5	410	2	AE1669	aminotripeptidase
133	32	60.4	467	2	AH3437	DNA repair protein	206	31	58.5	410	2	AD1297	probable ligase [i
134	32	60.4	470	2	G70566	probable membrane	207	31	58.5	411	2	B87200	Leu/ile/Val-bindin
135	32	60.4	480	2	H95376	hypothetical prote	208	31	58.5	415	2	AE3522	adenosylmethionine
136	32	60.4	488	2	T16195	hypothetical prote	209	31	58.5	450	1	B69835	DNA repair protein
137	32	60.4	491	2	A49993	glycylpeptide N-ter	210	31	58.5	455	1	A71251	unknown protein f2
138	32	60.4	492	2	C72417	sugar kinase, FGGY	211	31	58.5	467	2	G96837	conserved hypothet
139	32	60.4	500	2	D97221	xylose kinase [i	212	31	58.5	468	2	C82449	hypothetical prote
140	32	60.4	507	2	AG3187	aldehyde dehydroge	213	31	58.5	471	2	T04911	hypothetical prote
141	32	60.4	526	2	A75581	flavin monooamine o	214	31	58.5	485	2	H64164	hypothetical prote
142	32	60.4	547	2	T30269	hypothetical prote	215	31	58.5	486	1	A57601	transcription fact
143	32	60.4	588	2	C95252	L-fucose isomerase	216	31	58.5	497	1	S33938	penton protein [II
144	32	60.4	588	2	A99717	L-fucose isomerase	217	31	58.5	502	2	T25669	hypothetical prote
145	32	60.4	591	2	T25636	hypothetical prote	218	31	58.5	515	2	I39073	interferon alpha-b
146	32	60.4	604	2	F64081	isomerase fuci (EC	219	31	58.5	530	2	F90893	probable kinase [i
147	32	60.4	605	2	S25151	invertase, cell-wa	220	31	58.5	530	2	C85724	probable kinase yd
148	32	60.4	630	2	A70117	probable long-chain	221	31	58.5	530	2	B64905	sugar kinase homol
149	32	60.4	699	2	D82798	phage-related term	222	31	58.5	534	2	T52649	beta-1,2-xylosyltr
150	32	60.4	705	2	C84406	hypothetical prote	223	31	58.5	537	2	G82640	DNA methyltransfer
151	32	60.4	747	2	G87402	hypothetical prote	224	31	58.5	542	2	E70815	hypothetical prote
152	32	60.4	814	2	T00953	hypothetical prote	225	31	58.5	546	2	B37491	major capsid prote
153	32	60.4	827	1	S10639	fructose phosphotr	226	31	58.5	549	2	AB0987	alpha,alpha-trehal
154	32	60.4	859	2	C87358	hypothetical prote	227	31	58.5	549	2	H87178	probable DNA helic
155	32	60.4	956	2	S76194	hypothetical prote	228	31	58.5	577	2	A61372	hemolysin, extrace
156	32	60.4	1129	2	T25635	hypothetical prote	229	31	58.5	602	2	AF3490	DNA-directed DNA p
157	32	60.4	1175	2	T25634	hypothetical prote	230	31	58.5	612	2	F82429	protein-export mem
158	32	60.4	1192	2	T33157	hypothetical prote	231	31	58.5	614	2	T06741	hypothetical prote
159	32	60.4	1374	2	AE3259	extracellular seri	232	31	58.5	625	2	E70411	general secretion
160	32	60.4	1772	2	T36105	probable large gly	233	31	58.5	648	1	H69878	probable protein k
161	32	60.4	1791	2	T02909	hypothetical prote	234	31	58.5	650	2	C69678	involved in polyke
162	32	60.4	2082	2	T37056	probable multi-dom	235	31	58.5	678	2	A87698	NAD(+) synthetase,
163	32	60.4	3190	2	T13828	CREB-binding prote	236	31	58.5	696	2	AI2849	GGDEF family prote
164	31	58.5	57	2	A95869	hypothetical prote	237	31	58.5	696	2	G77626	hypothetical prote
165	31	58.5	76	2	AG2908	hypothetical prote	238	31	58.5	707	2	A46302	PRB-associated spl
166	31	58.5	131	2	AG0120	probable lipoprote	239	31	58.5	768	2	T00073	hypothetical prote
167	31	58.5	140	2	A70552	hypothetical prote	240	31	58.5	805	1	A55054	calpain (EC 3.4.22
168	31	58.5	157	2	A72558	hypothetical prote	241	31	58.5	821	2	E95245	penicillin-binding
169	31	58.5	159	2	D87586	hypothetical prote	242	31	58.5	821	2	B99110	peptidoglycan glyc
170	31	58.5	163	2	T27978	hypothetical prote	243	31	58.5	844	2	F86231	hypothetical prote
171	31	58.5	212	2	B75109	hypothetical prote	244	31	58.5	847	1	IJHUBD	desmocolin 3b pre
172	31	58.5	223	2	A70642	probable rplD prot	245	31	58.5	883	2	AI2618	DNA mismatch repair
173	31	58.5	226	2	AF2529	ThiJ family protei	246	31	58.5	894	2	S54786	multifunctional be
174	31	58.5	230	2	B86646	hypothetical prote	247	31	58.5	901	1	IJHUBA	desmocolin 3a pre
175	31	58.5	238	2	T02118	hypothetical prote	248	31	58.5	904	2	A97401	DNA mismatch repair



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:01:12 ; Search time 20 Seconds

(without alignments)

48.096 Million cell updates/sec

Title: US-10-697-055-7

Perfect score: 53

Sequence: 1 TPRVTGGAM 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	561	1 WMBE65	65K lower matrix p
2	53	100.0	561	1 WMBE65	65K lower matrix p
3	38	71.7	383	2 C96581	hypothetical prote
4	37	69.8	653	2 C87280	penicillin-binding
5	36	67.9	85	2 H95303	hypothetical prote
6	36	67.9	254	2 S25281	glTF protein precu
7	36	67.9	354	2 T46683	probable lipopate p
8	36	67.9	388	2 T28693	hypothetical prote
9	36	67.9	404	2 T02294	hypothetical prote
10	36	67.9	788	2 C70441	lipopate-protein li
11	36	67.9	799	2 T01663	1,4-alpha-glucan b
12	36	67.9	799	2 T02981	1,4-alpha-glucan b
13	35	66.0	223	2 AH2893	conserved hypothet
14	35	66.0	304	2 C97669	hypothetical prote
15	35	66.0	336	2 D72243	rod shape-determin
16	35	66.0	345	2 A53138	gas1 homolog - hum
17	35	66.0	359	2 JC7280	cytokine receptor-
18	35	66.0	522	2 T18583	glucosylceramidase
19	35	66.0	524	2 H86228	hypothetical prote
20	35	66.0	539	2 S57972	hypothetical prote
21	35	66.0	923	2 E70820	hypothetical glyci
22	35	66.0	3643	2 T36410	probable polyketid
23	34	64.2	211	2 T49671	hypothetical prote
24	34	64.2	247	2 S64065	hypothetical prote
25	34	64.2	252	2 H83822	hypothetical prote
26	34	64.2	267	2 A84233	hypothetical prote
27	34	64.2	269	2 AE1114	hydroxyethylthiaz
28	34	64.2	269	2 AF1475	hydroxyethylthiaz
29	34	64.2	288	2 T04622	prohibitin-like pr

30	34	64.2	298	2 C75430	uricase - Deinococ
31	34	64.2	370	2 AH2629	2'-deoxycytidine 5
32	34	64.2	370	2 E97411	hypothetical prote
33	34	64.2	380	1 VCVGCN	coat protein - cuc
34	34	64.2	380	2 P95101	aminotransferase,
35	34	64.2	380	2 G97969	hypothetical prote
36	34	64.2	408	2 G70925	probable PPE prote
37	34	64.2	409	2 G90481	dehydrogenase, pro
38	34	64.2	440	2 G96600	protein Fl4J16.24
39	34	64.2	940	2 A38149	RNA-directed RNA p
40	34	64.2	1844	2 T51890	related to Nup98-N
41	34	64.2	3300	2 D70575	probable PPE prote
42	33	62.3	23	2 S47208	T-cell receptor J-
43	33	62.3	111	2 B37266	Ig kappa chain V r
44	33	62.3	111	2 I38740	PL7-6 antibody lig
45	33	62.3	113	2 JC2270	hypothetical prote
46	33	62.3	125	2 G84604	probable acetyltra
47	33	62.3	147	2 T36374	hypothetical prote
48	33	62.3	154	2 S55017	hypothetical 16.7K
49	33	62.3	164	2 JQ1252	hypothetical prote
50	33	62.3	180	2 AD3294	hypothetical prote
51	33	62.3	213	2 C86575	hypothetical prote
52	33	62.3	213	2 B72049	hypothetical prote
53	33	62.3	219	2 JC7563	neurite outgrowth
54	33	62.3	223	1 QBE26	BZLF2 protein - hu
55	33	62.3	225	2 S55720	ribosomal protein
56	33	62.3	249	2 B75110	lipopate-protein li
57	33	62.3	249	2 B71024	probable lipopate p
58	33	62.3	249	2 T00169	anti repressor - S
59	33	62.3	250	2 H89989	anti repressor (im
60	33	62.3	260	2 T32525	hypothetical prote
61	33	62.3	311	2 S06619	syndecan-1 precurs
62	33	62.3	314	2 F96527	protein F27J15.20
63	33	62.3	321	2 JC5460	intracellular alka
64	33	62.3	322	2 G83922	lipopate-protein li
65	33	62.3	333	2 G86632	hypothetical prote
66	33	62.3	345	2 T34031	hypothetical prote
67	33	62.3	347	2 T48610	hypothetical prote
68	33	62.3	386	2 T12048	ribosomal protein
69	33	62.3	396	1 R5XL1A	ribosomal protein
70	33	62.3	396	1 R5XL1B	ribosomal protein
71	33	62.3	404	2 C70403	ferredoxin oxidore
72	33	62.3	407	1 R5PFL1	ribosomal protein
73	33	62.3	407	2 T48308	60S ribosomal prot
74	33	62.3	421	1 JC4277	ribosomal protein
75	33	62.3	426	1 T09551	ribosomal protein
76	33	62.3	444	2 D26421	shufion B' - Esch
77	33	62.3	445	2 D81716	hypothetical prote
78	33	62.3	485	2 F95238	PTS system, membra
79	33	62.3	488	2 T19854	hypothetical prote
80	33	62.3	508	2 G98102	hypothetical prote
81	33	62.3	560	2 A38731	alpha-1A adrenergi
82	33	62.3	706	2 T05565	hypothetical prote
83	33	62.3	706	2 G72499	hypothetical prote
84	33	62.3	799	2 T48889	serine/threonine p
85	33	62.3	838	2 T05410	hypothetical prote
86	33	62.3	1018	2 T13525	hypothetical prote
87	33	62.3	1036	2 A29832	HPI layer surface
88	33	62.3	1093	2 I38533	AF17 protein - hum
89	33	62.3	1244	2 S25327	cytoskeleton assem
90	32	60.4	143	2 S67619	ribosomal protein
91	32	60.4	154	2 E87307	hypothetical prote
92	32	60.4	175	2 A70952	hypothetical prote
93	32	60.4	196	1 S55332	macrophage infecti
94	32	60.4	217	2 G70324	hypothetical prote
95	32	60.4	222	2 B87380	hypothetical prote
96	32	60.4	230	2 H65084	hypothetical prote
97	32	60.4	235	2 A86150	TiN6.24 protein -
98	32	60.4	239	2 E75445	hypothetical prote
99	32	60.4	242	2 T02473	hypothetical prote
100	32	60.4	242	2 G84315	cobalt transport p
101	32	60.4	247	2 A75254	conserved hypothet
102	32	60.4	258	2 AD1365	anti-repressor [Ba

us-10-697-055-7.rsp

Sun Mar 7 13:59:04 2004

FT MOD RES 1 ACETYLATION (BY SIMILARITY).  
 SQ SEQUENCE 142 AA; 15756 MW; 5459FD6B08EEAE21 CRC64;  
 Query Match 60.4%; Score 32; DB 1; Length 142;  
 Best local Similarity 100.0%; Pred. No. 53;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RVTGGG 8  
 |||||  
 Db 67 RVTGGG 72

Search completed: March 7, 2004, 13:03:18  
 Job time : 26 secs

RESULT 49  
SLA1 YEAST STANDARD; PRT; 1244 AA.  
AC P32750;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytoskeleton assembly control protein SLA1.  
GN SLA1 OR YBL007C OR YBL0321.  
OS Saccharomycetes cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DDY 228;  
RX MEDLINE=93328765; PubMed=8335689;  
RA Holtzman D.A., Yang S., Drubin D.G.;  
RT "Synthetic-lethal interactions identify two novel genes, SLA1 and  
SLA2, that control membrane cytoskeleton assembly in Saccharomycetes  
cerevisiae.";  
RL J. Cell Biol. 122:635-644 (1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=93070613; PubMed=1441753;  
RA Delaveau T., Jacq C., Perea J.;  
RT "Sequence of a 12.7 kb segment of yeast chromosome II identifies a  
PDR-like gene and several new open reading frames.";  
RL Yeast 8:761-768 (1992).  
RN [3]  
RP INTERACTION WITH LBS5.  
RX MEDLINE=22287476; PubMed=12398763;  
RA Dewar H., Warren D.F., Gardiner F.C., Gourlay C.G., Satish N.,  
Richardson M.R., Andrews P.D., Ayscough K.R.;  
RT "Novel proteins linking the actin cytoskeleton to the endocytic  
machinery in Saccharomycetes cerevisiae.";  
RL Mol. Biol. Cell 13:3646-3661 (2002).  
CC [-] FUNCTION: Essential for the proper formation of the cortical  
actin cytoskeleton. Involved in controlling the size of cortical  
patches perhaps by regulating the nucleation of filaments at the  
cortex.  
CC [-] SUBUNIT: Interacts with LBS5.  
CC [-] SIMILARITY: Contains 3 SH3 domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; Z22810; CAA80463.1; -  
CC EMBL; Z35768; CAA84826.1; -  
CC EMBL; S47695; AAB23985.1; -  
CC PIR; S25327; S25327.  
CC HSSP; Q03526; IAWJ.  
CC GeneOnline; I38442; -  
CC SGD; S0000103; SLA1.  
CC GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.  
CC GO; GO:0008092; P:cytoskeletal protein binding; IMP.  
CC GO; GO:0007015; P:actin filament organization; IMP.  
CC GO; GO:0007047; P:cell wall organization and biogenesis; IMP.  
CC GO; GO:0006897; P:enocytosis; IMP.  
CC GO; GO:0007121; P:polar budding; IMP.  
CC InterPro; IPR001452; SH3.  
CC InterPro; IPR007131; SHD1.  
CC Pfam; PF00018; SH3; 3.  
CC Pfam; PF03983; SHD1; 1.  
CC PRINTS; PR00452; SH3DOMAIN.  
CC ProDom; PD000066; SH3; 3.  
CC SMART; SM00326; SH3; 3.

DR KW PROSITE; PS50002; SH3; 3.  
FT Cytoskeleton; Actin-binding; SH3 domain; Repeat.  
FT SH3 1.  
FT DOMAIN 8 69  
FT SH3 2.  
FT SH3 3.  
FT SH3 3.  
FT SH3 3.  
FT 16 X 7 AA APPROXIMATE REPEATS OF  
FT T-G-G-A-M-M-P.  
FT 1.  
FT REPEAT 868 874  
FT REPEAT 877 883  
FT REPEAT 887 893  
FT REPEAT 923 929  
FT REPEAT 923 929  
FT REPEAT 945 951  
FT REPEAT 1003 1009  
FT REPEAT 1020 1026  
FT REPEAT 1031 1037  
FT REPEAT 1048 1054  
FT REPEAT 1065 1071  
FT REPEAT 1084 1090  
FT REPEAT 1129 1135  
FT REPEAT 1155 1161  
FT REPEAT 1170 1176  
FT REPEAT 1185 1191  
FT REPEAT 1200 1206  
FT REPEAT 1244 AA; 135848 MW; 7FD85AA776407624 CRC64;  
SQ SEQUENCE  
Query Match 62.3%; Score 33; DB 1; Length 1244;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
Db 1062 PQTGGGAM 1070  
-----  
RESULT 50  
RS16 SACEX  
ID RS16 SACEX STANDARD; PRT; 142 AA.  
AC Q876B4; Q876B5;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 40S ribosomal protein S16.  
GN RPS16.  
OS Saccharomycetes exiguus (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=34358;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS 379;  
RC MEDLINE=22482865; PubMed=12594514;  
RC Langkjaer R.B., Clifton P.F., Johnston M., Piskur J.;  
RA "Yeast genome duplication was followed by asynchronous differentiation  
of duplicated genes.";  
RT Nature 421:848-852 (2003).  
RL Nature 421:848-852 (2003).  
CC [-] SIMILARITY: Belongs to the S9P family of ribosomal proteins.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AY144901; AAO32465.1; -  
CC EMBL; AY144900; AAO32464.1; -  
CC InterPro; IPR000754; Ribosomal\_S9.  
CC Pfam; PF00380; Ribosomal\_S9; 1.  
CC ProDom; PD001627; Ribosomal\_S9; 1.  
CC PROSITE; PS00360; RIBOSOMAL\_S9; 1.  
KW Ribosomal protein; Acetylation.  
FT INIT\_MET 0 0 BY SIMILARITY.

RESULT 47  
HP12 DEIRA  
ID HP12 DEIRA STANDARD; PRT; 1036 AA.  
AC P13126; 1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Hexagonally packed intermediate-layer surface protein precursor.  
GN HPI  
OS Deinococcus radiodurans.  
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=Saik;  
RX MEDLINE=88032846; PubMed=3667529;  
RA Peters J., Peters M., Lottspeich F., Schaefer W., Baumeister W.;  
RT "Nucleotide sequence analysis of the gene encoding the Deinococcus  
RT radiodurans surface protein, derived amino acid sequence, and  
RT complementary protein chemical studies."  
RL J. Bacteriol. 169:5216-5223(1987).  
CC -!- FUNCTION: Shape maintenance, possible protection from noxious  
CC enzymes or exogenous and unsettling DNA, and may mediate homotypic  
CC cell-cell contacts.  
CC -!- SUBCELLULAR LOCATION: Cell wall. This bacterium is covered by a S-  
CC layer with hexagonal symmetry.  
CC -!- PTM: CONTAINS TIGHTLY BOUND REDUCING SUGARS (SIX PER POLYPEPTIDE  
CC CHAIN) AND FATTY ACIDS (COVALENTLY BOUND AND LOCATED IN THE N-  
CC TERMINAL REGION).  
CC -!- PTM: The N-terminus is blocked.  
CC -!- MISCELLANEOUS: THE HYDROPHILIC C-TERMINAL REGION RICH IN AROMATIC  
CC AA COULD BE ENGAGED IN INTERACTIONS WITH NUCLEIC ACIDS, AND THE  
CC BOUND FATTY ACIDS AND THE N-TERMINAL REGION COULD SERVE TO ANCHOR  
CC THE LAYER TO THE OUTER MEMBRANE OF D. RADIOURANS. HPI LAYER  
CC CONTAIN ABOUT 30% BETA STRUCTURE AND VIRTUALLY NO ALPHA HELIX.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M17895; AAA23335.1; -  
CC PIR; A29832; A29832.  
CC InterPro; IPR000437; Prok lipoprot S.  
CC PROSITE; PS00013; PROKAR\_LIPOPROTEIN; Signal; Cell wall; S-layer.  
KW Glycoprotein; lipoprotein; Signal; Cell wall; S-layer.  
FT SIGNAL 1 17  
FT CHAIN 18 1036  
FT FT 74 86  
FT FT 256 275  
FT FT 642 754  
FT FT 60 250  
FT FT 985 1030  
FT FT 1036 AA; 108028 MW; AAF98206A74AEE CRC64;  
SQ SEQUENCE 1036 AA; 108028 MW; 108028 MW; 108028 MW; 108028 MW; 108028 MW;  
Query Match 62.3%; Score 33; DB 1; Length 1036;  
Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 TPRVTGGGAM 10  
Db 879 TPNVTGNSAL 888  
RESULT 48  
AF17\_HUMAN

ID AF17 HUMAN STANDARD; PRT; 1093 AA.  
AC P55138;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE AF-17 protein.  
GN MLLT6 OR AF17.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94336695; PubMed=8058765;  
RA Prasad R., Leshkowitz D., Gu Y., Alder H., Nakamura T., Saito H.,  
RA Huebner K., Berger R., Croce C.M., Canaani E.;  
RT "Leucine-zipper dimerization motif encoded by the AF17 gene fused to  
RT ALL-1 (MLL) in acute leukemia."  
RL Proc. Natl. Acad. Sci. U.S.A. 91:8107-8111(1994).  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- DISEASE: Involved in acute leukemias through a chromosomal  
CC translocation t(11;17)(q23;q21) that involves MLL/HRX and MLLT6.  
CC The result is a rogue activator protein.  
CC -!- SIMILARITY: Contains 2 PHD-type zinc fingers.  
CC -!- SIMILARITY: HIGH, TO AF10.  
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chromocancer/Genes/AF17.html".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U07932; AAA21145.1; -  
CC PIR; I38533; I38533.  
CC Genew; HGNC:7138; MLLT6.  
CC MIM; 600328; -  
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.  
CC InterPro; IPR001965; znf\_PHD.  
CC Pfam; PF00628; PHD; 1.  
CC SMART; SM00249; PHD; 2.  
CC PROSITE; PS01359; ZF\_PHD\_1; 1.  
CC PROSITE; PS00016; ZF\_PHD\_2; 2.  
KW Nuclear protein; Zinc-finger; Repeat; Chromosomal translocation;  
KW Proto-oncogene.  
FT ZN\_FING 5 57  
FT ZN\_FING 51 68  
FT ZN\_FING 117 180  
FT DOMAIN 190 211  
FT DOMAIN 275 282  
FT DOMAIN 326 338  
FT DOMAIN 729 764  
FT DOMAIN 822 829  
FT DOMAIN 834 862  
FT DOMAIN 935 984  
FT DOMAIN 1040 1051  
FT DOMAIN 1069 1080  
FT SITE 551 551  
FT SITE  
SQ SEQUENCE 1093 AA; 112021 MW; F60042A6D3BF579E CRC64;  
Query Match 62.3%; Score 33; DB 1; Length 1093;  
Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 TPRVTGGGA 9  
Db 429 SPHTVGGGA 437

DR InterPro; IPR008271; Ser thr\_pkin\_AS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00564; PQQ; 4.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation.  
 FT DOMAIN 16 272 PROTEIN KINASE  
 FT NP\_BIND 22 30 ATP (BY SIMILARITY).  
 FT BINDING 44 44 ATP (BY SIMILARITY).  
 FT ACT\_SITE 138 138 BY SIMILARITY.  
 SQ SEQUENCE 807 AA; 85231 MW; 66C274219155D091 CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 807;  
 Best Local Similarity 85.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 DB 655 PAVTGGG 661  
 RESULT 46  
 PNKL ARATH STANDARD; PRT; 870 AA.  
 ID Q8L5Y9; Q49372; O49374;  
 DC 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable pantothenate kinase 1 (EC 2.7.1.33) (Pantothenic acid kinase  
 1).  
 GN AT4G32180 OR F10M6.180 OR F10N7.10.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eumids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,  
 RA Kreis M., Deiseny M., Puigdomenech P., Watson M., Schmidheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
 RA Weitzengger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Moelmann P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,  
 RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker R., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maare A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzl A.,  
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Muehlen A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chedford F., Cooke R., Berger C., Monfort A., Casasuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Baigues M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,  
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,  
 RA Stoneking P., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,  
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
 RA Du H., Ali J., Berghoff A., Jones K., Dron K., Cotton M., Joshi C.,  
 RA Antonou B., Zidan M., Strong C., Sun H., Lamar B., Yordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
 RA Granat S., Shohdy N., Hasegawa A., Haneed A., Lodhi M., Johnson A.,  
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 402:769-777(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Anisari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayaishizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 genome.";  
 RL Science 302:842-846(2003).  
 CC -!- FUNCTION: Plays a role in the physiological regulation of the  
 intracellular CoA concentration (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + pantothenate = ADP + D-4'-  
 phosphopantothenate.  
 CC -!- ENZYME REGULATION: Regulated by feedback inhibition by malonyl-CoA  
 (By similarity).  
 CC -!- PATHWAY: Coenzyme A (CoA) biosynthesis; first step.  
 CC -!- SIMILARITY: Belongs to the eukaryotic pantothenate kinase family.  
 CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous  
 gene model prediction.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL021811; CAA16972.1; ALT\_SEQ.  
 CC EMBL; AL161580; CAB79936.1; ALT\_SEQ.  
 CC EMBL; AL021636; CAA16571.1; ALT\_SEQ.  
 CC EMBL; AY099839; AAM20690.1; --  
 CC InterPro; IPR002791; DUF89.  
 CC InterPro; IPR004567; Pank\_eukar.  
 CC Pfam; PF01937; DUF89; 1.  
 CC Pfam; PF03630; Fumble; 1.  
 CC TIGRfam; TIGR00555; pank\_eukar; 1.  
 KW Transferrase; Kinase; ATP-binding; Coenzyme A biosynthesis.  
 SQ SEQUENCE 870 AA; 96234 MW; 23CECA42A3D8EA6C CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 870;  
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RVTGGGA 9  
 DB 144 KVTGGGA 150

SQ SEQUENCE 692 AA; 76745 MW; 0E58310C4F3EFD80 CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 692;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
 DB 513 TPNETGGG 520

RESULT 44  
 AFSK STRCO STANDARD; PRT; 799 AA.  
 AC P54741; Q9P365; Q9L002;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine protein kinase afsk (EC 2.7.1.37).  
 GN AFSK OR SC04423 OR SC6F11.21 OR SC6D6.01.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=94341568; PubMed=8063104;  
 RA Matsumoto A., Hong S.K., Ishizuka H., Horinouchi S., Beppu T.;  
 RT "Phosphorylation of the AfsR protein involved in secondary metabolism  
 in Streptomyces species by a eukaryotic-type protein kinase.";  
 RL Gene 146:47-56(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=96186909; PubMed=8635757;  
 RA Ueda K., Uneyama T., Beppu T., Horinouchi S.;  
 RT "The aerial mycelium-defective phenotype of Streptomyces griseus  
 resulting from A-factor deficiency is suppressed by a Ser/Thr kinase  
 of S. coelicolor A3(2).";  
 RL Gene 169:91-95(1996).  
 RN [3]  
 RP REVISIONS TO 239-240.  
 RA Matsumoto A., Hong S., Ishizuka H., Horinouchi S., Beppu T.,  
 RA Uneyama T.,  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieger T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajadream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM BY  
 PHOSPHORYLATING, ON BOTH SER AND THR. THE AFSR GLOBAL REGULATORY  
 PROTEIN INVOLVED IN THE CONTROL OF SECONDARY METABOLISM.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- PTM: Autophosphorylated on serine and threonine residues.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D45382; BAA08229.2; -;  
 DR EMBL; AL939120; CAD55483.1; -;  
 DR PhosSite; P54741; -;  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SMO0564; EQQ; 9.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation; Complete proteome.  
 FT DOMAIN 16 271 PROTEIN KINASE  
 FT NP\_BIND 22 30 ATP (BY SIMILARITY).  
 FT BINDING 44 44 ATP (BY SIMILARITY).  
 FT ACT\_SITE 138 138 BY SIMILARITY.  
 SQ SEQUENCE 799 AA; 83787 MW; 4BE9BED4169F6F5B CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 799;  
 Best Local Similarity 85.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
 DB 648 PAVTGGG 654

RESULT 45  
 AFSK STRGR STANDARD; PRT; 807 AA.  
 AC P54742;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine protein kinase afsk (EC 2.7.1.37).  
 GN AFSK  
 OS Streptomyces griseus.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96186909; PubMed=8635757;  
 RA Ueda K., Uneyama T., Beppu T., Horinouchi S.;  
 RT "The aerial mycelium-defective phenotype of Streptomyces griseus  
 resulting from A-factor deficiency is suppressed by a Ser/Thr kinase  
 of S. coelicolor A3(2).";  
 RL Gene 169:91-95(1996).  
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM BY  
 PHOSPHORYLATING, ON BOTH SER AND THR. THE AFSR GLOBAL REGULATORY  
 PROTEIN INVOLVED IN THE CONTROL OF SECONDARY METABOLISM (BY  
 SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- PTM: Autophosphorylated on serine and threonine residues (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D45246; BAA08203.1; -;  
 DR PhosSite; P54742; -;  
 DR InterPro; IPR000719; Prot\_kinase.

FT DOMAIN 116 127 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 128 153 2 (POTENTIAL).  
FT DOMAIN 154 186 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 164 186 3 (POTENTIAL).  
FT DOMAIN 187 207 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 208 232 4 (POTENTIAL).  
FT DOMAIN 233 245 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 246 269 5 (POTENTIAL).  
FT DOMAIN 270 342 6 (POTENTIAL).  
FT TRANSMEM 343 367 7 (POTENTIAL).  
FT DOMAIN 368 374 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 375 399 8 (POTENTIAL).  
FT DOMAIN 400 561 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT LIPID 413 413 S-palmitoyl cysteine (potential).  
FT DOMAIN 21 57 POLY-RICH.  
FT DOMAIN 416 421 POLY-ARG.  
FT CONFLICT 424 483 AVGHWRRASTGDARSDCAPSPRIAPPGLAPLTAHPGAGS  
ADTPTQDSVSSRRKPASA -> SLRPPPLASLDRRAFLRL  
POPSHRSPPGSPSPHCTPCGGLGRHAGDAGFLQSKASLR  
(IN REF. 1).  
SQ SEQUENCE 561 AA; 59354 MW; B6537DCADAF7BE27 CRC64;  
Query Match 62.3%; Score 33; DB 1; Length 561;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PRVTGGAM 10  
Db 43 PGATGGGAV 51  
RESULT 42  
ID PD11 HUMAN STANDARD; PRT; 663 AA.  
AC Q9ULC6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein-arginine deiminase type I (EC 3.5.3.15) (Peptidylarginine  
deiminase I).  
GN PAD11 OR PD11.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Epidermis;  
RA Ishigami A., Guerrin-Weber M., Sebbag M., Serre G., Senahu T.;  
RT "Molecular cloning of peptidylarginine deiminase type I cDNA from a  
human epidermis.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Catalyzes the deimination of arginine residues of  
proteins (By similarity).  
CC -1- CATALYTIC ACTIVITY: Protein L-arginine + H(2)O = protein L-  
citrulline + NH(3).  
CC -1- COFACTOR: Requires calcium ions (By similarity).  
CC -1- SIMILARITY: Belongs to the protein arginine deiminase family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AB033768; BAA85771.1; -.  
CC Genew; HGNC:18367; PAD11.  
CC MIM; 607934; -.  
CC InterPro; IPR004303; Prot\_arg\_deim.

DR Pfam; PF03068; PAD; 1.  
KW Hydrolase; Calcium-binding; Multigene family.  
FT CA\_BIND 506 517 EF-HAND (POTENTIAL).  
SQ SEQUENCE 663 AA; 74607 MW; 0BD8D460634BE2D9 CRC64;  
Query Match 62.3%; Score 33; DB 1; Length 663;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TPRTVGGG 8  
Db 415 SPPTVGGG 422  
RESULT 43  
ID EFG\_STAEP STANDARD; PRT; 692 AA.  
AC Q8C082;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Elongation factor G (EF-G).  
GN FUS1 OR SE0311.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 12228;  
RX PubMed=12950922;  
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;  
RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
Staphylococcus epidermidis strain (ATCC 12228).";  
RT Mol. Microbiol. 49:1577-1593 (2003).  
RL -1- FUNCTION: This protein promotes the GTP-dependent translocation of  
the nascent protein chain from the A-site to the P-site of the  
ribosome.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
EF-G/EF-2 subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AE016745; AA003908.1; -.  
CC HAMAP; MF\_00054; -; 1.  
DR InterPro; IPR000795; EF\_GTPbind.  
DR InterPro; IPR000640; EFG\_C.  
DR InterPro; IPR009022; EFG\_III\_V.  
DR InterPro; IPR005517; EFG\_IV.  
DR InterPro; IPR004161; EFTU\_D2.  
DR InterPro; IPR005225; Small\_GTP.  
DR InterPro; IPR009000; Translat\_factor.  
DR Pfam; PF00679; EFG\_C; 1.  
DR Pfam; PF03764; EFG\_IV; 1.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR PRINTS; PR00315; ELONGATNFCT.  
DR TIGRFS; TIGR00231; small\_GTP; 1.  
DR PROSITE; PS00301; EFATOR\_GTP; 1.  
DR Elongation factor; Protein biosynthesis; GTP-binding;  
KW Complete proteome.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT NP\_BIND 16 23 GTP (BY SIMILARITY).  
FT NP\_BIND 80 84 GTP (BY SIMILARITY).  
FT NP\_BIND 134 137 GTP (BY SIMILARITY).  
FT



RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foslaker A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of *Drosophila melanogaster*."  
RT Science 287:2195-2195(2000).  
RL [3]  
RN  
RP SEQUENCE OF 1-89 FROM N.A., AND TISSUE SPECIFICITY.  
RC STRAIN=Oregon-R;  
RX MEDLINE=96217926; PubMed=8675012;  
RA Merli C., Bergstrom D.E., Cygan J.A., Blackman R.K.;  
RT "Promoter specificity mediates the independent regulation of  
neighboring genes.";  
RL Genes Dev. 10:1260-1270(1996).  
CC -I- TISSUE SPECIFICITY: Vital for proper neuronal development and hatching.  
CC -I- TISSUE SPECIFICITY: Embryonic expression is in small clusters of  
cells along posterior margin of most segments, brain and  
segmentally repeating pattern along midline of nerve cord.  
CC Expressed in embryonic, larval and adult gonads of both sexes, and  
larval imaginal disks.  
CC -I- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically  
throughout development with highest levels in pupae and adults.  
CC -I- MISCELLANEOUS: Readthrough of the terminator UGA may occur between  
the codons for 332-Glu and 334-Arg.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L31349; AAC37219.2; -;  
CC EMBL; AE003583; AAF51246.1; ALT\_TERM.  
CC FLYBase; FBgn0011818; oaf.  
CC GO; GO:0016321; P:female meiosis chromosome segregation; IMP.  
CC Developmental protein.  
CC CHAIN 1 487 OUT AT FIRST PROTEIN.  
CC FT CHAIN 1 332 OUT AT FIRST SHORT PROTEIN.  
CC FT DOMAIN 450 480 PRO/SER/THR-RICH.  
CC FT DOMAIN 392 396 POLY-THR.  
CC FT DOMAIN 402 405 POLY-LEU.  
CC FT DOMAIN 456 459 POLY-SER.  
CC FT DOMAIN 472 480 POLY-SER.  
CC FT DOMAIN 391 391 A -> G (IN REF. 1).  
CC FT CONFLICT 391 391 A -> G (IN REF. 1).  
CC FT CONFLICT 400 400 R -> H (IN REF. 1).

FT CONFLICT 413 413 A -> T (IN REF. 1).  
SQ SEQUENCE 487 AA; 53728 MW; F0D7A117BE358B3F CRC64;  
Query Match 62.3%; Score 33; DB 1; Length 487;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TPRVTGGGA 9  
DB 373 TATITGGGA 381  
RESULT 41  
ID ALAD RAT STANDARD; PRT; 561 AA.  
AC P23944;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Alpha-1D adrenergic receptor (Alpha 1D-adrenoceptor) (Alpha-1A  
DE adrenergic receptor) (RA42).  
GN ADRA1D OR ADRA1A.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain cortex;  
RX MEDLINE=91177889; PubMed=1706716;  
RA Lomasney J.W., Cotecchia S., Lorenz W., Leung W.-Y., Schwinn D.A.,  
RA Yang-Feng T.L., Brownstein M., Lefkowitz R.J., Caron M.G.;  
RT "Molecular cloning and expression of the cDNA for the alpha 1A-  
RT adrenergic receptor. The gene for which is located on human  
chromosome 5.";  
RT J. Biol. Chem. 266:6365-6369(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95114877; PubMed=7815325;  
RA Schwinn D.A., Johnston G.I., Page S.O., Mosley M.J., Wilson K.H.,  
RA Worman N.P., Campbell S., Fidock M.D., Furness L.M.,  
RA Parry-Smith D.J., Peter B., Bailey D.S.;  
RT "Cloning and pharmacological characterization of human alpha-1  
RT adrenergic receptors: sequence corrections and direct comparison with  
other species homologues.";  
RT J. Pharmacol. Exp. Ther. 272:134-142(1995).  
CC -I- FUNCTION: This alpha-adrenergic receptor mediates its effect  
through the influx of extracellular calcium.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -I- TISSUE SPECIFICITY: Vas deferens, hippocampus, cerebral cortex,  
aorta, brain stem, heart and spleen.  
CC -I- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M60654; AAA63477.1; -;  
CC EMBL; L31771; AAB59704.1; -;  
CC PIR; A38731; A38731.  
CC InterPro; IPR000276; GPCR\_Rhodpsn.  
CC Pfam; PF00001; 7tm.1; 1.  
CC PRINTS; PR00237; GPCR\_Rhodopsin.  
CC PROSITE; PS00237; G-PROTEIN\_RECEP\_FL\_1; 1.  
CC PROSITE; PS0262; G-PROTEIN\_RECEP\_FL\_2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
CC Multigene family; Phosphorylation; Lipoprotein; Palmitate.  
CC DOMAIN 1 90 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 91 115 1 (POTENTIAL).  
CC FT



RL Curr. Biol. 7:977-986(1997).  
 RP [3]  
 RX SEQUENCE FROM N.A.  
 RA MEDLINE=20184068; PubMed=10717241;  
 RA Senba S., Ouyang H., Han S.-Y., Kato Y., Horii A.;  
 RT "Analysis of the candidate target genes for mutation in microsatellite  
 RT instability-positive cancers of the colorectum, stomach, and  
 RT endometrium";  
 RL Int. J. Oncol. 16:731-737(2000).  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,  
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
 RA Schackwitz W.S., Shewerwood J.K., Witral L.A., Nickerson D.A.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RL [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow, and Muscle;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: INVOLVED IN CELL CYCLE ARREST WHEN DNA DAMAGE HAS  
 CC OCCURRED OR WHEN UNLIGATED DNA IS PRESENT. BINDS TO AND  
 CC PHOSPHORYLATES CDC25A, CDC25B AND CDC25C. PHOSPHORYLATION OF  
 CC CDC25C CREATES A BINDING SITE FOR 14-3-3 PROTEIN WHICH INHIBITS  
 CC CDC25C. THIS PREVENTS ACTIVATION OF THE CDC2-CYCLIN B COMPLEX AND  
 CC PREVENTS MITOTIC ENTRY. PHOSPHORYLATES CDC25C ON SERINE-216.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Expressed ubiquitously with the most abundant  
 CC expression in thymus, testis, small intestine and colon.  
 CC -!- PTM: Phosphorylated.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1  
 CC subfamily.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF016582; AAC51736.1; -;  
 DR EMBL; AF032874; AAB88852.1; -;  
 DR EMBL; AB032387; BAA84577.1; -;  
 DR EMBL; AB032377; BAA84577.1; JOINED.  
 DR EMBL; AB032378; BAA84577.1; JOINED.  
 DR EMBL; AB032379; BAA84577.1; JOINED.  
 DR EMBL; AB032380; BAA84577.1; JOINED.  
 DR EMBL; AB032381; BAA84577.1; JOINED.  
 DR EMBL; AB032382; BAA84577.1; JOINED.  
 DR EMBL; AB032383; BAA84577.1; JOINED.  
 DR EMBL; AB032384; BAA84577.1; JOINED.  
 DR EMBL; AB032385; BAA84577.1; JOINED.  
 DR EMBL; AB032386; BAA84577.1; JOINED.

DR EMBL; AF527555; AAM78553.1; -;  
 DR EMBL; BC004202; AAH04202.1; -;  
 DR EMBL; BC017575; AAH17575.1; -;  
 DR PDB; 1IA8; 1A-APR-01.  
 DR Genew; HGNC:1925; CHEK1.  
 DR GK; O14757; -;  
 DR MIM; 603078; -;  
 DR GO; GO:0005711; C:meiotic chromosome; TAS.  
 DR GO; GO:0004672; P:protein kinase activity; TAS.  
 DR GO; GO:0000077; P:DNA damage response, signal transduction re.; TAS.  
 DR GO; GO:0007276; P:gametogenesis; TAS.  
 DR GO; GO:0007131; P:meiotic recombination; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO; GO:0000079; P:regulation of CDK activity; TAS.  
 DR InterPro; IPR000719; Ser Thr Kinase.  
 DR InterPro; IPR008271; Ser Thr pkin AS.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Prot Kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;  
 KW DNA damage; Nuclear protein; Phosphorylation; 3D-structure.  
 FT DOMAIN 9 265 PROTEIN KINASE.  
 FT NP BIND 15 23 ATP (BY SIMILARITY).  
 FT BINDING 38 38 ATP (BY SIMILARITY).  
 FT ACT SITE 130 130 BY SIMILARITY.  
 FT SEQUENCE 476 AA; 54419 MW; 0ABD0FAB67F49F67 CRC64;  
 SQ  
 Query Match 62.3%; Score 33; DB 1; Length 476;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTCGG 8  
 DB 276 PRVTCGG 282  
 RESULT 40  
 ID\_DROME STANDARD; PRT; 487 AA.  
 AC Q9NLA6; Q24556;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Out at first protein [Contains: Out at first short protein].  
 GN OAF OR CG9884.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 RX MEDLINE=95286050; PubMed=7768442;  
 RA Bergstrom D.E., Merli C.A., Cygan J.A., Cygan R.K.;  
 RT "Regulatory autonomy and molecular characterization of the Drosophila  
 RT out at first gene";  
 RL Genetics 139:1331-1346(1995).  
 RL [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

```

DR MM: 60555; -: C: integral to membrane; NAS.
DR GO: GO:0016021; -: C: integral to membrane; NAS.
DR GO: GO:0004252; F: serine-type endopeptidase activity; NAS.
DR GO: GO:0006508; P: proteolysis and peptidolysis; NAS.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR002172; LDL_receptor A.
DR InterPro: IPR001254; Peptidase S1.
DR InterPro: IPR001314; Peptidase S1A.
DR InterPro: IPR001190; Scrv_receptor.
DR Pfam: PF00057; ldl_recept_a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00192; LDLa; 1.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS01209; LDLA_1; FALSE_NEG.
DR PROSITE: PS00068; LDLA_2; FALSE_NEG.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS00287; SRCR_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Transmembrane; Signal-anchor.
KW CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 32
FT TRANSMEM 33 53
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT LDL-RECEPTOR CLASS A.
FT SRCR.
FT SERINE PROTEASE.
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 245 245
FT ACT_SITE 290 290
FT ACT_SITE 367 367
FT SITE 204 205
FT DISULFID 64 83
FT DISULFID 77 92
FT DISULFID 127 183
FT DISULFID 140 193
FT DISULFID 196 310
FT DISULFID 230 246
FT DISULFID 356 372
FT DISULFID 383 410
FT CARBOHYD 130 130
FT CARBOHYD 178 178
FT CONFLICT 1 31
FT SEQUENCE 437 AA; 48204 MW; 351B2FD4A8657B12 CRC64;
Query Match 62.3%; Score 33; DB 1; Length 437;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGG 7
Db 202 TPRVVG 208

RESULT 38
SHU4_ECOLI STANDARD; PRT; 444 AA.
AC P09748;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Shufflon protein B'.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN RP
RP SEQUENCE FROM N.A.
RC PLASMID=IncII R64;

QY 1 TPRVTGG 7
Db 202 TPRVVG 208

RESULT 39
CHK1_HUMAN STANDARD; PRT; 476 AA.
AC Q14757;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Chk1 (EC 2.7.1.-).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN RP
RP SEQUENCE FROM N.A.
RX MEDLINE=97426625; PubMed=9278511;
RA Sanchez Y., Wong C., Thoma R.S., Richman R., Wu Z., Piwnicka-Worms H.,
RA Elledge S.J.;
RT "Conservation of the Chk1 checkpoint pathway in mammals: linkage of
RT DNA damage to Cdk regulation through Cdc25."
RL Science 277:1497-1501 (1997).
RN RP
RP SEQUENCE FROM N.A.
RX MEDLINE=98044285; PubMed=9382850;
RA Flagg G., Plug A.W., Dunks K.M., Mundt K.E., Ford J.C.,
RA Quiggle M.R.E., Taylor E.M., Westphal C.H., Ashley T., Hoekstra M.F.,
RA Carr A.M.;
RT "Atm-dependent interactions of a mammalian chk1 homolog with meiotic
RT chromosomes."

```



Sun Mar 7 13:59:04 2004

RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=94092742; PubMed=8268230;  
 RA Bagni C., Mariottini P., Annesi F., Amaldi F.;  
 RT "Human ribosomal protein L4: cloning and sequencing of the cDNA and  
 RL primary structure of the protein.";  
 RN Biochim. Biophys. Acta 1216:475-478 (1993).  
 [2]  
 RP REVISIONS.  
 RA Bagni C.;  
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymphoma;  
 RA Kato S.;  
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21864036; PubMed=11875025;  
 RA Yoshihama M., Uechi T., Asakawa S., Kawasaki K., Kato S., Higa S.,  
 RT Maeda N., Minoshima S., Tanaka T., Shimizu N., Kameuchi N.;  
 RL "The human ribosomal protein genes: sequencing and comparative  
 analysis of 73 genes.";  
 RN Genome Res. 12:379-390 (2002).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon, Eye, Muscle, Pancreas, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [6]  
 RP -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X73974; CAA52154.1; -  
 CC EMBL; L20868; AAA60281.1; ALT\_SEQ.  
 CC EMBL; D33660; BAA04887.1; -  
 CC EMBL; BC001365; AAH01365.1; -  
 CC EMBL; BC005817; AAH05817.1; -  
 CC EMBL; BC007259; AAH07259.1; -  
 CC EMBL; BC007748; AAH07748.1; -  
 CC EMBL; BC007996; AAH07996.1; -  
 CC EMBL; BC009888; AAH09888.1; -  
 CC EMBL; BC010151; AAH10151.1; -  
 CC EMBL; BC014653; AAH14653.1; -  
 CC EMBL; AB061820; BAB79458.1; -  
 CC PIR; T09551; -  
 CC SWISS-2DPAGE; P36578; HUMAN.  
 CC Genew; HGNC:10353; RPL4.

DR GK, P36578; -;  
 DR MIM; 180479; -;  
 DR GO:0005842; C:cytosolic large ribosomal subunit (sensu Eu. . .; TAS.  
 DR GO:0003723; F:RNA binding; TAS.  
 DR GO:0003735; F:structural constituent of ribosome; TAS.  
 DR GO:0006412; P:protein biosynthesis; TAS.  
 DR InterPro; IPR002136; Ribosomal\_L4/L1E.  
 DR Pfam; PF00573; Ribosomal\_L4; 1.  
 DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.  
 KW Ribosomal protein.  
 FT DOMAIN 364 427 LYS-RICH.  
 FT CONFLICT 3 3 C -> V (IN REF. 1).  
 FT CONFLICT 36 36 I -> M (IN REF. 1).  
 FT CONFLICT 63 63 S -> R (IN REF. 1).  
 FT CONFLICT 147 147 V -> F (IN REF. 1).  
 FT CONFLICT 201 201 MISSING (IN REF. 1).  
 SQ SEQUENCE 427 AA; 47697 MW; 4785ED31699CD792 CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 427;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 DB 77 PRVRGGG 83  
 RESULT 36  
 TMS4\_MOUSE  
 ID TMS4\_MOUSE STANDARD; PRT; 435 AA.  
 AC Q8VCA5.  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transmembrane protease, serine 4 (BC 3.4.21.-) (Channel-activating  
 DE protease 2) (mCAP2).  
 GN TMRPSS4 OR CAP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22144321; PubMed=12149280;  
 RA Vuagniaux G., Vallet V., Jager N.F., Hummler E., Rossier B.C.;  
 RT "Synergistic activation of ENAC by three membrane-bound channel-  
 RT activating serine proteases (mCAP1, mCAP2, and mCAP3) and serum-  
 RT glucocorticoid-regulated kinase (Sgk1) in Xenopus oocytes.";  
 RL J. Gen. Physiol. 120:191-201 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF134732; RAD32206.1; --  
 DR InterPro; IPR002136; Ribosomal\_L4/L1e.  
 DR Pfam; PF00573; Ribosomal\_L4.1  
 DR PROSITE; PS00399; RIBOSOMAL\_L1e; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 408 AA; 44869 MW; DAD8C75E8CDA7594 CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 408;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 DB 84 PRVTGGG 90  
 RESULT 31  
 SEPR THESE STANDARD; PRT; 408 AA.  
 ID SEPR THESE STANDARD; PRT; 408 AA.  
 AC P80146; 1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Extracellular serine proteinase precursor (EC 3.4.21.-).  
 OS Thermus sp. (strain Rt41A).  
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
 OC Thermus.  
 OX NCBI\_TaxID=32063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96004470; PubMed=7551039;  
 RA Munro G.K., McHale R.H., Saul D.J., Reeves R.A., Bergquist P.L.;  
 RT "A gene encoding a thermophilic alkaline serine proteinase from  
 RT Thermus sp. strain Rt41A and its expression in Escherichia coli."  
 RL Microbiology 141:1731-1738(1995).  
 RN [2]  
 RP SEQUENCE OF 131-145, AND CHARACTERIZATION.  
 RX MEDLINE=92362605; PubMed=1499549;  
 RA Peek K., Daniel R.M., Monk C., Parker L., Coolbear T.;  
 RT "Purification and characterization of a thermostable proteinase  
 RT isolated from Thermus sp. strain Rt41A."  
 RL Eur. J. Biochem. 207:1035-1044(1992).  
 CC -!- FUNCTION: THERMOSTABLE SERINE PROTEINASE WITH PREFERRED  
 CC ACTIVITY FOR AMINO ACIDS WITH AROMATIC SIDE GROUPS AT THE  
 CC P1, SIDE OF THE SCISSILE BOND.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PTM: Contains 4 Cys residues that form two disulfide bonds.  
 CC -!- PTM: This proteinase has a 0.7% carbohydrate content.  
 CC -!- SIMILARITY: Belongs to peptidase family S8.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U17342; AAA82980.1; --  
 DR HSPSP; Q99405; IMPT.  
 DR MEROPS; S08.008; --  
 DR InterPro; IPR000209; Peptidase S8.  
 DR InterPro; IPR000437; Prok lipoprot\_S.  
 DR Pfam; PF00882; Peptidase\_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASE ASP; 1.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Signal.  
 FT SIGNAL 1 19  
 FT PROPEP 20 130  
 FT CHAIN 131 408  
 FT ACT SITE 169 169  
 FT ACT SITE 202 202  
 FT ACT SITE 354 354  
 FT SEQUENCE 408 AA; 42284 MW; 7F48716F541A6CA7 CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 408;  
 Best Local Similarity 60.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 357 TPRVTGGGAM 366  
 RESULT 32  
 RL4 MOUSE STANDARD; PRT; 419 AA.  
 ID RL4 MOUSE STANDARD; PRT; 419 AA.  
 AC Q9DBE6; Q9CY08;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 60S ribosomal protein L4 (L1).  
 GN RPL4  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryonic liver, and Small intestine;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,  
 RA Arakawa T., Hara A., Fukumishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Caasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayaishizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RN Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scherz T.E.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scherz T.E.,  
 RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,



Leonard S., Meyer R., Mulvaney E., Ozeraky P., Riley A., Strommatt C., Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N., Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J., Kirchhoff K., Toth K., King L., Wilson R.K., Murphy G., Bancroft I., Martienssen R., McCombie W.R., Walsby R.K., Stiekema W., Pohl T., Voelckert G., Wambut R., Duesterhoeft A., Stiekema W., Pohl T., Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S., Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W., Rameberger U., Wedler H., Balke K., Wedler E., Peters S., van Staveren M., Dirks W., Mooijman P., Klein Iankhorst R., Weitzenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S., Feldpausch M., Lamberth S., Villarroel R., Giesen J., Ardiles W., Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H., Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F., "Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana." Nature 408:823-826(2000).

[2] SEQUENCE OF 4-142 FROM N.A.

RP STRAIN=CV. Columbia;

RC STRAIN=CV. Columbia;

RA Desprez T., Anselm J., Chiappello H., Rouze P., Caboche M., Hofte H.; Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.

[3] SEQUENCE OF 108-191 FROM N.A.

RP STRAIN=CV. Columbia;

RC Berthomieu P., Guerrier D., Giraudat J.; Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE OF 197-407 FROM N.A.

RP STRAIN=CV. Columbia;

RC Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.; Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.

RL -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; A1462973; CAB6041.1; -

DR EMBL; Z18118; CAA79104.1; -

DR EMBL; Z18460; CAA79191.1; -

DR EMBL; Z17589; CAA79004.1; -

PIR; T48308; T48308.

DR InterPro; IPR002136; Ribosomal L4/L1E.

DR Pfam; PF00573; Ribosomal L4; 1.

DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.

KW Ribosomal protein.

FT CONFLICT 181 G -> V (IN REF. 3).

FT CONFLICT 362 E -> D (IN REF. 4).

SQ SEQUENCE 407 AA; 44722 MW; 026FC2852B7A2038 CRC64;

Query Match 62.3%; Score 33; DB 1; Length 407;

Best Local Similarity 85.7%; Pred. No. 1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 1;

QY 2 PRVTGGG 8

DB 83 PRVTGGG 89

RESULT 29

RL4 DROME STANDARD; PRT; 407 AA.

ID - RL4 DROME

AC P09180;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE 60S ribosomal protein L4 (L1).

GN RPL4 OR RPL1.

Drosophila melanogaster (Fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R;

RX MEDLINE=89098414; PubMed=2492096;

RA Rafti F., Gargiulo G., Manzi A., Malva C., Graziani F.; "Sequence of the ribosomal protein cDNA of D. melanogaster homologous to the L1 ribosomal protein gene of X. laevis." Nucleic Acids Res. 17:456-456(1989).

RL Nucleic Acids Res. 17:456-456(1989).

RN [2]

RP SEQUENCE OF 63-186 FROM N.A.

RX MEDLINE=89262486; PubMed=3133637;

RA Rafti F., Gargiulo G., Manzi A., Malva C., Grossi G., Andone S., Graziani F.; "Isolation and structural analysis of a ribosomal protein gene in D. melanogaster." Nucleic Acids Res. 16:4915-4926(1988).

RL Nucleic Acids Res. 16:4915-4926(1988).

CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X13382; CAA31759.1; -

DR EMBL; X06881; CAA29998.1; -

DR PIR; S02209; RSFFL1.

DR FlyBase; FBgn003279; Rpl1.

DR InterPro; IPR002136; Ribosomal L4/L1E.

DR Pfam; PF00573; Ribosomal L4; 1.

DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.

KW Ribosomal protein.

FT CONFLICT 117 K -> R (IN REF. 2).

SQ SEQUENCE 407 AA; 45755 MW; F35D789870B043C CRC64;

Query Match 62.3%; Score 33; DB 1; Length 407;

Best Local Similarity 85.7%; Pred. No. 1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 1;

QY 2 PRVTGGG 8

DB 80 PRVTGGG 86

RESULT 30

RL4 PRUAR STANDARD; PRT; 408 AA.

ID - RL4 PRUAR

AC Q9X97;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 60S ribosomal protein L4 (L1).

GN RPL4.

OS Prunus armeniaca (Apricot).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid1; Rosales; Rosaceae; Amygdaloideae; Prunus.

OX NCBI\_TaxID=36596;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. Bergeron; TISSUE=Mesocarp;

RA Mbeguie-A-Mbeguie D., Fils-Lycaon B.R.; "Molecular cloning and nucleotide sequence of a 60S ribosomal protein L1 from apricot (Prunus armeniaca cv. Bergeron)." Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RL -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

```

RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwälder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Lauzo C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.P., Cottet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Ligouri R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Millescher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shimp S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana."
RL Nature 408:820-822(2000).
CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AC016661; AAF23293.1; -
CC InterPro: IPR002136; Ribosomal_L4/L1E.
CC Pfam: PF00573; Ribosomal_L4; 1.
CC PROSITE: PS00939; RIBOSOMAL_L1E; 1.
CC Ribosomal protein.
CC
CC Query Match 62.3%; Score 33; DB 1; Length 406;
CC Best Local Similarity 85.7%; Pred. No. 1e+02;
CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 PRVTGGG 8
DB 82 PRVPGGG 88

```

RESULT 28

RL4A ARATH STANDARD; PRT; 407 AA.

AC P49691; O9LYZ8;

DT 01-FEB-1996 (Rel. 33, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-FEB-2004 (Rel. 43, Last annotation update)

BE 60S ribosomal protein L4-1 (L1).

GN RPL4A OR RPL4 OR RPL1 OR AT5G02870 OR F9G14\_180.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=21016721; PubMed=11130714;

RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,

RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,

RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,

RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,

RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,

RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,

RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,

RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,

RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,

RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,

```

RX MEDLINE=86135987; PubMed=3841512;
RA Lorenzi F., Ruberti I., Bozzoni I., Pierandrei-Amaldi P., Amaldi F.,
RA "Nucleotide sequence of the L1 ribosomal protein gene of Xenopus
RT laevis: remarkable sequence homology among introns."
RL EMBO J. 4:3483-3488(1985).
CC [2]
CC SEQUENCE OF 286-396 FROM N.A.
CC MEDLINE=82262793; PubMed=7049839;
CC Amaldi F., Beccari E., Bozzoni I., Luo Z.-X., Pierandrei-Amaldi P.;
CC "Nucleotide sequences of cloned cDNA fragments specific for six
RT Xenopus laevis ribosomal proteins."
RL Gene 17:311-316(1982).
CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X05217; CAA28844.1; -
CC EMBL: V01438; CAA24699.1; -
CC PIR: B24579; RSXLB.
CC InterPro: IPR002136; Ribosomal_L4/L1E.
CC Pfam: PF00573; Ribosomal_L4; 1.
CC PROSITE: PS00939; RIBOSOMAL_L1E; 1.
CC Ribosomal protein.
CC
CC NON_TER 1 1 TDLT -> LICA (IN REF. 2).
CC CONFLICT 286 289 L -> P (IN REF. 2).
CC CONFLICT 321 321 L -> P (IN REF. 2).
CC CONFLICT 337 339 HAI -> MH (IN REF. 2).
CC CONFLICT 382 382 R -> K (IN REF. 2).
CC SEQUENCE 396 AA; 44906 MW; BFC1E2D604572FF0 CRC64;
QY 2 PRVTGGG 8
DB 76 PRVPGGG 82

```

Query Match 62.3%; Score 33; DB 1; Length 396;

Best Local Similarity 85.7%; Pred. No. 98;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8

DB 76 PRVPGGG 82

RESULT 27

RL4A ARATH STANDARD; PRT; 406 AA.

AC Q9SF40;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 60S ribosomal protein L4-2 (L1).

GN RPL4B OR AT3G09630 OR F11F8.22.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=21016720; PubMed=11130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,

RA Partmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,

RA Delsen M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,

RA De Simone V., Choisine N., Artiguenave F., Robert C., Brottier P.,

RA Wincker P., Catolico L., Weisenbach J., Saurin W., Quetier F.,

RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,

RA Wurmbach E., Drzonek H., Brfhe H., Jordan N., Bangert S.,

RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,

RA Verzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,



OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Fulton R., Woldmann P.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF000196; AAC24253.1; -.

DR PIR; T34031; T34031.

DR WormPep; B0041.4; CE07669.

DR InterPro; IPR002136; Ribosomal\_L4/L1E.

DR Pfam; PF00573; Ribosomal\_L4; 1.

DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.

KW Ribosomal protein.

SQ SEQUENCE 345 AA; 38659 MW; C798B9DB0CED60AD CRC64;

Query Match 62.3%; Score 33; DB 1; Length 345;

Best Local Similarity 85.7%; Pred. No. 85;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRVTGGG 8

Db 76 PRVRGGG 82

RESULT 24

RL4\_URECA

ID RL4\_URECA STANDARD; PRT; 386 AA.

AC P49165;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE 60S ribosomal protein L4 (L1).

GN RPL4 OR RPL1.

OS Urechis caupo (Innkeeper worm) (Spoonworm).

OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.

OX NCBI\_TaxID=6431;

RN [1]

RP SEQUENCE FROM N.A.

RA Rosenthal E.T.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U30495; AAA74021.1; -.

DR PIR; T12048; T12048.

DR InterPro; IPR002136; Ribosomal\_L4/L1E.

DR Pfam; PF00573; Ribosomal\_L4; 1.

DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.

KW Ribosomal protein.

SQ SEQUENCE 386 AA; 43135 MW; 23811EBEF015DD99 CRC64;

Query Match 62.3%; Score 33; DB 1; Length 386;

Best Local Similarity 85.7%; Pred. No. 96;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRVTGGG 8

Db 76 PRVRGGG 82

RESULT 25

RL4A\_XENLA

ID RL4A\_XENLA STANDARD; PRT; 396 AA.

AC P08429; O91843;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 60S ribosomal protein L4A (L1A).

GN RPL4A OR RPL1A.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RE MEDLINE=86135987; PubMed=3841512;

RA Loreni F., Ruberti I., Bozzoni I., Pierandrei-Amaldi P., Amaldi F.;

RT "Nucleotide sequence of the L1 ribosomal protein gene of Xenopus

laevis: remarkable sequence homology among introns.";

RL EMBJ. 4:3483-3488(1985).

CC -1- SIMILARITY: Belongs to the L4E family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X05216; CAA28843.1; -.

DR EMBL; X06552; CAA29796.1; -.

DR PIR; A24579; RSXL1A.

DR InterPro; IPR002136; Ribosomal\_L4/L1E.

DR Pfam; PF00573; Ribosomal\_L4; 1.

DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.

KW Ribosomal protein.

FT CONFLICT 246

FT CONFLICT 392

FT CONFLICT 392 P -> A (IN REF. 1; CAA29796).

SQ SEQUENCE 396 AA; 44935 MW; 2DA0FB693A1186BC CRC64;

Query Match 62.3%; Score 33; DB 1; Length 396;

Best Local Similarity 85.7%; Pred. No. 98;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 PRVTGGG 8

Db 81 PRVRGGG 87

RESULT 26

RL4B\_XENLA

ID RL4B\_XENLA STANDARD; PRT; 396 AA.

AC P02385;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE 60S ribosomal protein L4B (L1B) (Fragment).

GN RPL4B OR RPL1B.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

QY 2 PRVTGGGA 9  
 DB 240 PRVTGGAA 247

RESULT 22  
 SDC1\_MOUSE  
 ID SDC1\_MOUSE STANDARD; PRT; 311 AA.  
 AC P1828; Q62278; Q9WD22;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Syndecan-1 precursor (SYND1).  
 GN SDC1 OR SYND1 OR SYND-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Mammary gland;  
 RC MEDLINE=89174988; PubMed=2494134;  
 RA Saunders S., Jalkanen M., O'Farrell S., Bernfield M.;  
 RT "Molecular cloning of syndecan, an integral membrane proteoglycan.";  
 RL J. Cell Biol. 108:1547-1556 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=93352511; PubMed=8349612;  
 RA Vihtinen T., Auvinen P., Alanen-Kurki L., Jalkanen M.;  
 RT "Structural organization and genomic sequence of mouse syndecan-1  
 gene.";  
 RL J. Biol. Chem. 268:17261-17269 (1993).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=NMRI; TISSUE=Embryo;  
 RX MEDLINE=99303636; PubMed=10373479;  
 RA Romaris M., Coomans C., Ceulemans H., Bruystens A.-M., Vekemans S.,  
 RA David G.;  
 RT "Molecular polymorphism of the syndecans. Identification of a hypo-  
 glycanated murine syndecan-1 splice variant.";  
 RL J. Biol. Chem. 274:18667-18674 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=Cecch II; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Carninci P., Prange C.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Small D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP SEQUENCE OF 1-22 FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=93266605; PubMed=8496192;  
 RA Hinkes M.T., Goldberger O., Neumann P., Kokenyeji R., Bernfield M.;  
 RT "Organization and promoter activity of the mouse syndecan-1 gene.";

J. Biol. Chem. 268:11440-11448 (1993).  
 -1- FUNCTION: Cell surface proteoglycan that bears both heparan sulfate and chondroitin sulfate and that links the cytoskeleton to the interstitial matrix.  
 -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 -1- ALTERNATIVE PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 Name=1;  
 IsoId=P1828-1; Sequence=VSP\_007542;  
 Note=Major isoform;  
 Name=2;  
 IsoId=P1828-2; Sequence=VSP\_007542;  
 Note=Minor isoform;  
 -1- SIMILARITY: Belongs to the syndecan proteoglycan family.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; X15487; CAA33514.1; -;  
 EMBL; Z22532; CAA80254.1; -;  
 EMBL; AF134897; AAD42345.1; -;  
 EMBL; BC010560; AAA10560.1; -;  
 EMBL; L11565; AAA40159.1; -;  
 PIR; S06619; S06619.  
 MGI; MGI:1349162; Sdc1.  
 InterPro; IPR003585; Neurexin-like.  
 InterPro; IPR001050; Syndecan.  
 Pfam; PF01034; Syndecan; 1.  
 SMART; SM00294; 4.1m; 1.  
 PROSITE; PS00964; SYNDECAN; 1.  
 Proteoglycan; Heparan sulfate; Transmembrane; Glycoprotein; Signal;  
 Alternative splicing.  
 SIGNAL 1 17  
 CHAIN 18 311  
 DOMAIN 18 252  
 TRANSMEM 253 277  
 DOMAIN 278 311  
 CARBOHYD 37 37  
 CARBOHYD 43 43  
 CARBOHYD 45 45  
 CARBOHYD 47 47  
 CARBOHYD 207 207  
 CARBOHYD 217 217  
 SITE 251 252  
 VARSPLIC 50 93  
 CONFLICT 21 21 L -> V (IN REF. 5).  
 SEQUENCE 311 AA; 32904 MW; 283FEC396FF40FCE CRC64;  
 Query Match 62.3%; Score 33; DB 1; Length 311;  
 Best Local Similarity 85.7%; Pred. No. 77;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGG 8  
 DB 179 PRVEGGG 185  
 RESULT 23  
 RL4\_CAEEL  
 ID RL4\_CAEEL STANDARD; PRT; 345 AA.  
 AC O02056;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 60S ribosomal protein L4.  
 GN RPL-4 OR R0041.4.  
 OS Caenorhabditis elegans.



```

CC CC -1- SIMILARITY: COAT PROTEINS FROM CARMOVIRUSES (CARMV, MNV, TCV),
CC CC DIANTHOVIRUSES (RCNV), SOBEMOVIRUSES (MCMV, SEMV), TOMBUSVIRUSES
CC CC (AMCV, CNV, CYRSV, AND TESV), AND TNV ARE RELATED IN THEIR S
CC CC REGION.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M25270; AAA42904.1; --
CC CC PIR; JAO131; VCVGCV.
CC CC HSP; P11795; 2TBV.
CC CC InterPro; IPR008975; Viral_cap_coat.
CC CC InterPro; IPR000937; Viral_coat.
CC CC Pfam; PF00729; Viral_coat.1.
CC CC PRINTS; PR00233; ICOSAHEDRAL.
CC CC PROSITE; PS00555; ICOSAH_VIR_COAT_S; 1.
CC CC -----
CC CC Coat protein. 1 92 R DOMAIN, INTERACTION WITH RNA.
CC CC DOMAIN 93 254 S DOMAIN, VIRION SHELL.
CC CC DOMAIN 255 380 P DOMAIN, PROJECTING.
CC CC SEQUENCE 380 AA; 40878 MW; 793CB6B05CEG6669 CRC64;
CC CC -----
CC CC Query Match 64.2%; Score 34; DB 1; Length 380;
CC CC Best Local Similarity 50.0%; Pred. No. 61;
CC CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CC CC -----
CC CC QY 1 TPRVTGGGAM 10
CC CC ||:||||:
CC CC 318 TPSIGGGTI 327
CC CC -----
CC CC RESULT 18
CC CC YS92 MYCTU
CC CC ID YS92 MYCTU STANDARD; PRT; 408 AA.
CC CC AC Q10813;
CC CC DT 01-OCT-1996 (Rel. 34, Created)
CC CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC CC DE Hypothetical PPE-family protein RV2892C/MT2959/Mb2916C.
CC CC GN RV2892C OR MT2959 OR MTCY274.23C OR MB2916C.
CC CC OS Mycobacterium tuberculosis, and
CC CC OS Mycobacterium bovis.
CC CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC CC OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC CC OX NCBI_TaxID=1773, 1765;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
CC CC RX MEDLINE=98295987; PubMed=9634230;
CC CC RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
CC CC RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
CC CC RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
CC CC RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
CC CC RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
CC CC RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
CC CC RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
CC CC RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
CC CC RT "Deciphering the biology of Mycobacterium tuberculosis from the
CC CC RT complete genome sequence."
CC CC RL Nature 393:537-544 (1998).
CC CC [2]
CC CC RN [2]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
CC CC RX MEDLINE=22206494; PubMed=12218036;
CC CC RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
CC CC RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
CC CC RA Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M., Salzberg S.L.,
CC CC RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

```

```

RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains.";
RN J. Bacteriol. 184:5479-5490 (2002).
[3]
RP SEQUENCE FROM N.A.
RC SPECIES=M bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Gronlin S., Lacroix C., Monsempé C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- SIMILARITY: Belongs to the mycobacterial PPE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z74024; CAA98377.1; -.
DR EMBL; AE007119; AAK47285.1; -.
DR EMBL; BX248344; CAD96603.1; -.
DR PIR; G70925; G70925.
DR TIGR; MT2959; -.
DR TubercuList; RV2892c; -.
DR InterPro; IPR00030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 56 76 POTENTIAL.
FT SEQUENCE 408 AA; 41469 MW; 3E3D1F20D7827199 CRC64;
SQ
Query Match 64.2%; Score 34; DB 1; Length 408;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRTVYGGGG 10
DB 292 SPSVAGGGAV 301

RESULT 19
YZL2 EBV
ID YZL2 EBV STANDARD; PRT; 223 AA.
AC P03205;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical BZLF2 protein.
GN BZLF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OC NCBI_TaxID=10377;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrell B.G.;
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RT Nature 310:207-211(1984).
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/

```

```

OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OK NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CLIP 11262 / Serovar 6a;
RC MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RA "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + 4-methyl-5-(2-hydroxyethyl)-thiazole =
CC ADP + 4-methyl-5-(2-phosphoethyl)-thiazole.
CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: Belongs to the Thz kinase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL596164; CAC95574.1; -.
CC PIR; AF1475; AF1475.
CC ListList; LINO0341; -.
CC HAMAP; MF_00228; -; 1.
CC InterPro; IPR000417; Hyethyz_kinase.
CC Pfam; PF02110; HK; 1.
CC PRINTS; PR01099; HYETHYZKNASE.
CC TIGRFAMs; TIGR00694; thm; 1.
CC Thiamine biosynthesis; Transferase; Kinase; ATP-binding; Magnesium;
KW Complete proteome.
FT ACT_SITE 194 194 BASE (BY SIMILARITY).
FT METAL 91 91 MAGNESIUM 1 (BY SIMILARITY).
FT METAL 123 123 MAGNESIUM 1 (BY SIMILARITY).
FT SEQUENCE 269 AA; 28166 MW; 7288ABFD23A760A1 CRC64;
SQ
Query Match 64.2%; Score 34; DB 1; Length 269;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 PRVTGGGAM 10
DB ||:||||:
187 PRITGSGCL 195
RESULT 16
THIM_LISMO
ID THIM_LISMO STANDARD; PRT; 269 AA.
AC Q8VAX6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hydroxyethylthiazole kinase (EC 2.7.1.50) (4-methyl-5-beta-
DE hydroxyethylthiazole kinase) (Thz kinase) (TH kinase).
GN THIM OR LMO0316.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OK NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RA "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + 4-methyl-5-(2-hydroxyethyl)-thiazole =
CC ADP + 4-methyl-5-(2-phosphoethyl)-thiazole.
CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: Belongs to the Thz kinase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL591974; CAD00843.1; -.
CC PIR; AE1114; AE1114.
CC ListList; LMO00316; -.
CC HAMAP; MF_00228; -; 1.
CC InterPro; IPR000417; Hyethyz_kinase.
CC Pfam; PF02110; HK; 1.
CC PRINTS; PR01099; HYETHYZKNASE.
CC TIGRFAMs; TIGR00694; thm; 1.
CC Thiamine biosynthesis; Transferase; Kinase; ATP-binding; Magnesium;
KW Complete proteome.
FT ACT_SITE 194 194 BASE (BY SIMILARITY).
FT METAL 91 91 MAGNESIUM 1 (BY SIMILARITY).
FT METAL 123 123 MAGNESIUM 1 (BY SIMILARITY).
FT SEQUENCE 269 AA; 28052 MW; 74A3D672176242C CRC64;
SQ
Query Match 64.2%; Score 34; DB 1; Length 269;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 PRVTGGGAM 10
DB ||:||||:
187 PRITGSGCL 195
RESULT 17
COAT_CNV
ID COAT_CNV STANDARD; PRT; 380 AA.
AC P15183;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Coat protein.
OS Cucurbit necrosis virus (CNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OC NCBI_TaxID=12143;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204896; PubMed=2705296;
RA Rochon D.M., Tremaine J.H.;
RA "Complete nucleotide sequence of the cucumber necrosis virus genome.";
RL Virology 169:251-259(1989).
CC -!- SUBUNIT: THE VIRUS COAT IS COMPOSED OF 180 COPIES OF THE COAT
CC PROTEIN.

```

MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shmolen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RN SEQUENCE OF 6-206 FROM N.A.  
RP TISSUE=Placenta;  
RC MEDLINE=98440284; PubMed=9767165;  
RX Robb L., Hartley L., Wang C.-C., Harvey R.P., Begley C.G.;  
RT "Musculin: a murine basic helix-loop-helix transcription factor gene  
RT expressed in embryonic skeletal muscle.";  
RL Mech. Dev. 76:197-201(1998).  
CC -!- FUNCTION: Transcription repressor capable of inhibiting the  
CC transactivation capability of E47. May play a role in regulating  
CC antigen-dependent B-cell differentiation.  
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another  
CC bHLH protein. Binds DNA as a homodimer or a heterodimer. Forms a  
CC heterodimer with E12/E47.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- TISSUE SPECIFICITY: Expressed in lymphoid tissues, B-cell lines  
CC and activated B cells.  
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
CC frameshift in position 189.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF060154; AAC15071.1; ALT\_FRAME.  
DR EMBL; BC006313; AAH06313.1; -  
DR EMBL; AF087036; AAC69870.1; -  
DR TRANSFAC; T04902; -  
DR Genew; HGNC:7321; MSC.  
DR MIM; 603628; -  
DR GO; GO:0005634; C:nucleus; TAS.  
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.  
DR InterPro; IPR001092; HLH\_basic.  
DR Pfam; PF00010; HLH; 1.  
DR SMART; SM00353; HLH; 1.  
DR PROSITE; PS00888; HLH; 1.  
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.  
FT DOMAIN 48 57  
FT DOMAIN 71 76  
FT DOMAIN 79 85  
FT DOMAIN 88 91  
FT DOMAIN 79 91  
FT DOMAIN 108 119  
FT DOMAIN 120 160  
FT CONFLICT 77 77  
V -> G (IN REF. 3).

FT CONFLICT 189 191 DSD -> ASA (IN REF. 1).  
SQ SEQUENCE 206 AA; 22067 MW; 21066E02553EB80C CRC64;  
Query Match 64.2%; Score 34; DB 1; Length 206;  
Best Local Similarity 85.7%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGG 8  
Db 75 PRVAGGG 81  
||| |||  
||| |||  
RESULT 14  
YGGI YEAST STANDARD; PRT; 247 AA.  
ID YGGI YEAST STANDARD; PRT; 247 AA.  
AC P53168;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 27.5 kDa protein in PYC1-UBC2 intergenic region.  
GN YGL061C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288C;  
RX MEDLINE=97377993; PubMed=9234674;  
RA Feuermann M., de Montigny J., Potier S., Souciet J.-L.;  
RT suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae  
RT chromosomes.";  
RL Yeast 13:861-869(1997).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; Z72583; CAA96764.1; -  
DR PIR; S64065; S64065.  
DR GermOnline; 141109; -  
DR SGD; S0003029; DUO1.  
DR GO; GO:0000778; C:condensed nuclear chromosome kinetochore; IPI.  
DR GO; GO:0005819; C:spindle; IPI.  
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IPI.  
DR GO; GO:0000071; P:mitotic spindle assembly (sensu Saccharomycetes); IPI.  
KW Hypothetical protein.  
SQ SEQUENCE 247 AA; 27473 MW; 787F8AF869E3C978 CRC64;  
Query Match 64.2%; Score 34; DB 1; Length 247;  
Best Local Similarity 75.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGA 9  
Db 35 PKSTGGGA 42  
RESULT 15  
THIM LISIN  
ID THIM LISIN STANDARD; PRT; 269 AA.  
AC Q9ZEW7; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hydroxyethylthiazole kinase (EC 2.7.1.50) (4-methyl-5-beta-  
DE hydroxyethylthiazole kinase) (Thz kinase) (TH kinase).  
GN THIM OR LIN0341.

RX MEDLINE=99059501; PubMed=9845368;  
 RA Ni M., Tepperman J.M., Quail P.H.;  
 RT "PIF3, a phytochrome B-interacting factor necessary for normal  
 RT photoinduced signal transduction, is a novel basic helix-loop-helix  
 RL protein.";  
 RL Cell 95:657-667(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Lee J., Yi H., Shin B., Song P.-S., Choi G.;  
 RT "Identification and characterization of three phytochrome-associated  
 RT proteins.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Schoenbohm C., Weishaar B.;  
 RT "Overview of the bHLH transcription factor gene family in Arabidopsis  
 RL thaliana.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Ienz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malti R., Marziali A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RL thaliana.";  
 RL Nature 408:816-820(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J., Kawai J.,  
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;  
 RT "Arabidopsis thaliana full-length cDNA.";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP FUNCTION.  
 RX MEDLINE=99394605; PubMed=10466729;  
 RA Ni M., Tepperman J.M., Quail P.H.;  
 RT "Binding of phytochrome B to its nuclear signalling partner PIF3 is  
 RT reversibly induced by light.";  
 RL Nature 400:781-784(1999).  
 RN [7]  
 RP FUNCTION.  
 RX MEDLINE=20259742; PubMed=10797009;  
 RA Martinez-Garcia J.F., Hug E., Quail P.H.;  
 RT "Direct targeting of light signals to a promoter element-bound  
 RT transcription factor.";  
 RL Science 288:859-863(2000).  
 RN [8]  
 RP GENE FAMILY ORGANIZATION, AND NOMENCLATURE.  
 RX MEDLINE=22598051; PubMed=12679534;  
 RA Heim M.A., Jakoby M., Werber M., Martin C., Weishaar B., Bailey P.C.;  
 RA "The basic helix-loop-helix transcription factor family in plants: a  
 RT genome-wide study of protein structure and functional diversity.";  
 RL Mol. Biol. Evol. 20:735-747(2003).  
 RL -i- FUNCTION: Transcription factor acting positively in the

CC phytochrome signaling pathway. Activates transcription by binding  
 CC to the G box (5'-CACGTG-3').  
 CC -i- SUBUNIT: Homodimer (probable). Can form a heterodimer with RFP1.  
 CC Phytochrome B binds specifically to DNA-bound PIF3, but only upon  
 CC red light induced conversion to the Pfr form (PfrB). Reconversion  
 CC to Pfr form causes rapid dissociation.  
 CC -i- SUBCELLULAR LOCATION: Nuclear.  
 CC -i- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF100156; AAC95156.1; -  
 CC EMBL; AF088280; AAC99771.1; -  
 CC EMBL; AF251693; AAL55715.1; -  
 CC EMBL; AC003970; AAC33213.1; -  
 CC EMBL; AK117255; BAC41930.1; -  
 CC PIR; H86228; H86228.  
 CC HGSP; P25912; IHLO.  
 CC TRANSFAC; T04492; -  
 CC InterPro; IPR001092; HLH\_basic.  
 CC Pfam; PF00010; HLH; 1.  
 CC SMART; SM00353; HLH; 1.  
 CC PROSITE; PS00888; HLH; 1.  
 CC Nuclear protein. Transcription regulation; DNA-binding;  
 KW Phytochrome signaling pathway.  
 FT DNA BIND 340 356 BASIC DOMAIN.  
 FT DOMAIN 357 393 HELIX-LOOP-HELIX MOTIF.  
 FT DOMAIN 353 358 POLY-ARG.  
 FT DOMAIN 422 428 POLY-ALA.  
 FT CONFLICT 15 15 E -> D (IN REF. 2).  
 FT CONFLICT 344 344 S -> L (IN REF. 2).  
 SQ SEQUENCE 524 AA; 56990 MW; 1044AC01D598DE7C CRC64;  
  
 Query Match 66.0%; Score 35; DB 1; Length 524;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2 PRVTGG 8  
 DB 470 PRVSGG 476  
  
 RESULT 13  
 ID MUSC\_HUMAN STANDARD; PRT; 206 AA.  
 AC O60682; O75946; Q9BRE7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Musculin (Activated B-cell factor-1) (ABF-1).  
 GN MSC OR ABFI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC TISSUE=B-cell;  
 RX MEDLINE=98252901; PubMed=9584154;  
 RA Maseari M.E., Rivera R.R., Voland J.R., Quong M.W., Breit T.M.,  
 RA van Dongen J.J.M., de Smit O., Murre C.;  
 RT "Characterization of ABF-1, a novel basic helix-loop-helix  
 RL transcription factor expressed in activated B lymphocytes.";  
 RN Mol. Cell. Biol. 18:3130-3139(1998).  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;



RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Mang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Kana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RL human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[6]  
RN SEQUENCE OF 234-359 FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner T., Glessi C., King B., Kochiwa H.,  
RA Fleschmann W., Gaasterland T., Nishida I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Wynshaw-Boris A., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690 (2001).  
CC -1- FUNCTION: Receptor for thymic stromal lymphopoietin (TSLP). Forms  
CC a functional complex with TSLP and IL7R which is capable of  
CC stimulating cell proliferation through activation of STAT3 and  
CC STAT5. Also activates JAK2. Implicated in the development of the  
CC hematopoietic system.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1 and  
CC isoform 3). Secreted (isoform 2).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=Q8CII9-1; Sequence=Displayed;  
CC Name=2; Synonyms=Soluble CRLM-2;  
CC IsoId=Q8CII9-2; Sequence=VSP\_008788, VSP\_008789;  
CC Name=3;  
CC IsoId=Q8CII9-3; Sequence=VSP\_008790;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: High level of expression in liver, lung and  
CC testis. Also expressed in heart, brain, spleen, thymus and bone  
CC marrow. Highly expressed in progenitors and myeloid cells. Isoform  
CC 2 is expressed in primary hematopoietic cells.  
CC -1- INDUCTION: Upregulated in the myeloid 32d cell line by granulocyte  
CC colony stimulating factor (G-CSF).  
CC -1- DOMAIN: The WSXWS motif appears to be necessary for proper protein  
CC folding and thereby efficient intracellular transport and cell-  
CC surface receptor binding.  
CC -1- DOMAIN: The BOX 1 motif is important for association with JAKs.  
CC -1- SIMILARITY: Belongs to the type I cytokine family of receptors.  
CC Subfamily 5.

CC -1- SIMILARITY: Contains 1 fibronectin type III domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AB019945; BAA92684.1; -  
CC EMBL; AB011333; BAA92159.1; -  
CC EMBL; AF232936; AAF61676.1; -  
CC EMBL; AF201963; AAF82189.1; -  
CC EMBL; BC023788; AAH23788.1; -  
CC EMBL; AK010291; BAB26827.1; -  
CC MGD; MGI:1889506; Tslpr.  
CC InterPro; IPR003961; FN\_III.  
CC Pfam; PF00044; fn3; 1.  
CC SMART; SM00060; FN3; 1.  
CC PROSITE; PS01355; HEMATOPO REC S\_F1; FALSE NEG.  
CC Receptor; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
KW SIGNAL 1 19  
FT CHAIN 20 359  
FT DOMAIN 20 232  
FT TRANSMEM 233 253  
FT DOMAIN 254 359  
FT DOMAIN 117 208  
FT SITE 201 205  
FT SITE 262 269  
FT DISULFID 68 82  
FT CARBOHYD 53 53  
FT CARBOHYD 122 122  
FT VARSPLIC 217 217  
FT VARSPLIC 218 359  
FT VARSPLIC 217 217  
FT VARSPLIC 217 217  
FT CONFLICT 15 15  
FT CONFLICT 51 51  
FT CONFLICT 87 87  
FT CONFLICT 179 179  
FT CONFLICT 309 309  
FT CONFLICT 359 AA; 37761 MW; F9C521C54B4AC9DD CRC64;  
SQ SEQUENCE 359 AA; 37761 MW; F9C521C54B4AC9DD CRC64;  
Query Match 66.0%; Score 35; DB 1; Length 359;  
Best Local Similarity 77.8%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
DB 333 PRGPGGGAM 341  
RESULT 12  
PIF3 ARATH  
ID PIF3 ARATH STANDARD; PRT; 524 AA.  
AC O80536; Q9SRC5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Phytocrome-interacting factor 3 (Phytochrome-associated protein 3)  
DE (Basic helix-loop-helix protein 8) (bHLH008).  
DE BHLH008 OR BHLH8 OR PIF3 OR PAP3 OR AT1G09530 OR F14J9.19.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OC NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;



```

DR InterPro: IPR001524; Glyco hydro 6.
DR Pfam: PF01341; Glyco hydro_6; 1.
DR PRINTS: PR00733; GLHYDRIASE6.
DR ProDom: PD003733; Glyco_hydro_6; 1.
DR PROSITE: PS00655; GLYCOSYL HYDROL F6_1; FALSE_NEG.
DR PROSITE: PS00656; GLYCOSYL HYDROL F6_2; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 27
FT CHAIN 28 321
FT ACT SITE 110 110
FT ACT SITE 149 149
FT ACT SITE 295 295
FT DISULFID 112 156
FT SEQUENCE 321 AA; 33694 MW; DC5E39EF3CCD114C CRC64;

Query Match 66.0%; Score 35; DB 1; Length 321;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 274 TPTRTGGGA 282

RESULT 10
GAS1_HUMAN
ID GAS1_HUMAN STANDARD; PRT; 345 AA.
AC P54826;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Growth-arrest-specific protein 1 precursor (GAS-1).
GN GAS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Liver;
RX MEDLINE=94173926; PubMed=8127893;
RA del Sal G., Collavin L., Ruaro M.E., Edomi P., Saccone S., Valle G.D.,
RA Schneider C.;
RT "Structure, function, and chromosome mapping of the
RT growth-suppressing human homologue of the murine gas1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1848-1852(1994).
CC -!- FUNCTION: Specific growth arrest protein involved in growth
CC suppression. Blocks entry to S phase. Prevents cycling of
CC normal and transformed cells.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; LI3698; AAA72368.1; -.
CC PIR; A53138; A53138.
CC Genew; HGNC:4165; GAS1.
CC MIM; 139185.
CC GO; GO:0046658; C:extrinsic to plasma membrane, GPI-anchored; ISS.
CC GO; GO:0007050; P:cell cycle arrest; TAS.
CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
CC GO; GO:0045749; P:negative regulation of S phase of mitotic c. . .; NAS.
CC Growth arrest; GPI-anchor; Signal; Glycoprotein; Lipoprotein.
FT SIGNAL 1 39
FT CHAIN 40 318
FT PROPEP 319 345
FT DOMAIN 87 91
POLY-ALA.

FT DOMAIN 285 288
FT DOMAIN 319 322
FT DOMAIN 337 341
FT LIPID 318 318
FT CARBOHYD 117 117
SQ SEQUENCE 345 AA; 35721 MW; 2AADS0FID3632F9D CRC64;

Query Match 66.0%; Score 35; DB 1; Length 345;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9
Db 148 PRVTGGGA 155

RESULT 11
CRL2_MOUSE
ID CRL2_MOUSE STANDARD; PRT; 359 AA.
AC Q8CII9; Q9CRJ6; Q9JIE7; Q9JIQ7; Q9JUH8; Q9JUD5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytokine receptor-like factor 2 precursor (Type I cytokine receptor
DE delta 1) (Cytokine receptor-like molecule-2) (CRLM-2) (Thymic stromal
DE lymphopoietin protein receptor) (TSLPR).
GN CRLF2 OR CRLM2 OR TSLPR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2).
RC TISSUE=Embryo;
RX MEDLINE=20329232; PubMed=10872831;
RA Hiroyama T., Iwana A., Morita Y., Nakamura Y., Shibuya A.,
RA Nakauchi H.;
RT "Molecular cloning and characterization of CRLM-2, a novel type I
RT cytokine receptor preferentially expressed in hematopoietic cells.";
RL Biochem. Biophys. Res. Commun. 272:224-229(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.
RC TISSUE=Lymphocytes;
RX MEDLINE=20197866; PubMed=10733486;
RA Fujio K., Nosaka T., Kojima T., Kawashima T., Yahata T.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K., Nishimura T.,
RA Kitamura T.;
RT "Molecular cloning of a novel type I cytokine receptor similar to the
RT common gamma chain.";
RL Blood 95:2204-2211(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6; TISSUE=Lymphocytes;
RX MEDLINE=20432254; PubMed=10974032;
RA Park L.S., Martin U., Garka K., Gliniak B., Di Santo J.P., Muller W.,
RA Largaespada D.A., Copeland N.G., Jenkins N.A., Farr A.G.,
RA Ziegler S.F., Morrissey P.J., Paxton R., Sims J.E.;
RT "Cloning of the murine thymic stromal lymphopoietin (TSLP) receptor.
RT Formation of a functional heteromeric complex requires interleukin 7
RT receptor.";
RL J. Exp. Med. 192:659-670(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=21177258; PubMed=10881176;
RA Pandey A., Ozaki K., Baumann H., Levin S.D., Puel A., Farr A.G.,
RA Ziegler S.F., Leonard W.J., Lodish H.F.;
RT "Cloning of a receptor subunit required for signaling by thymic
RT stromal lymphopoietin.";
RL Nat. Immunol. 1:59-64(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=FVB/N;
RX MEDLINE=22388257; PubMed=12477932;

```

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; A012186; A040186.1; -  
CC HAMAP; MF 00061; -; 1.  
CC InterPro; IPR006204; GHMP kinase.  
DR InterPro; IPR004424; IspE.  
DR Pfam; PF00288; GHMP kinases; 1.  
DR TIGRams; TIGR00154; ispe; 1. biosynthesis; ATP-binding;  
KW transferase; Kinase; Isoprene biosynthesis; ATP-binding;  
KW Complete proteome.  
FT NP BIND 106 116 ATP (POTENTIAL).  
SQ SEQUENCE 295 AA; 31085 MW; 18F318EEA2FAD384 CRC64;

Query Match 67.9%; Score 36; DB 1; Length 295;  
Best Local Similarity 66.7%; Pred. No. 20;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
|||:|:|:  
DB 242 TPRLTGSGS 250

RESULT 8  
GLGB MAIZE  
ID GLGB MAIZE STANDARD; PRT; 799 AA.  
AC Q08047;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 1,4-alpha-glucan branching enzyme IIB, Chloroplast precursor  
DE (EC 2.4.1.18) (Starch branching enzyme IIB) (Q-enzyme).  
GN SBE1

OS Zea mays (Maize)  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 58-65.  
RC STRAIN=CV. W64A X 192E; TISSUE=Endosperm;  
RX MEDLINE=94105320; PubMed=8278524;  
RA Fisher D.K., Boyer C.D., Hannah L.C.;  
RT "Starch branching enzyme II from maize endosperm.";  
RL Plant Physiol. 102:1045-1046(1993).  
RN [2]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 248-271 AND 305-315.  
RC STRAIN=CV. B73; TISSUE=Endosperm;  
RX MEDLINE=95152344; PubMed=7849565;  
RA Guan H.P., Baba T., Preiss J.;  
RT "Expression of branching enzyme II of maize endosperm in Escherichia  
coli.";  
RL Cell. Mol. Biol. 40:981-988(1994).

CC -!- FUNCTION: Catalyzes the formation of the alpha-1,6-glucosidic  
CC linkages in glycogen by scission of a 1,4-alpha-linked  
CC oligosaccharide from growing alpha-1,4-glucan chains and the  
CC subsequent attachment of the oligosaccharide to the alpha-1,6  
CC position.  
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of  
CC starch.

CC -!- PATHWAY: Starch biosynthesis; third step.  
CC -!- SUBUNIT: Monomer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Chloroplast; amyloplast.  
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; L08065; AAA18571.1; -  
CC PIR; T02981; T02981.  
CC MaizeDB; 63943; -  
CC InterPro; IPR006047; Alpha amyl cat.

DR InterPro; IPR004193; Glyco\_hydro\_13N.  
DR Pfam; PF00128; alpha-amylase; 1.  
DR Pfam; PF02922; isoamylase N; 1.  
KW Glycogen biosynthesis; Starch biosynthesis; Transferase;  
KW Glycosyltransferase; Amyloplast; Chloroplast; Transit peptide.  
FT TRANSIT 1 57 CHLOROPLAST (AMYLOPLAST).

FT CHAIN 58 799 1,4-ALPHA-GLUCAN BRANCHING ENZYME IIB.  
FT ACT SITE 341 341 BY SIMILARITY.  
FT ACT SITE 376 376 BY SIMILARITY.  
FT ACT SITE 381 381 BY SIMILARITY.  
FT ACT SITE 445 445 BY SIMILARITY.  
FT ACT SITE 447 447 BY SIMILARITY.  
FT ACT SITE 502 502 BY SIMILARITY.  
FT ACT SITE 569 569 BY SIMILARITY.  
FT ACT SITE 570 570 BY SIMILARITY.  
SQ SEQUENCE 799 AA; 90517 MW; 0B440E0377B8087A CRC64;

Query Match 67.9%; Score 36; DB 1; Length 799;  
Best Local Similarity 85.7%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
||:|:|:  
DB 17 PRLTGGG 23

RESULT 9  
GUNI STRHA  
ID GUNI STRHA STANDARD; PRT; 321 AA.  
AC P33682;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Endoglucanase I precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)  
DE (Cellulase I) (CMCase I) (CELL1).  
GN CELA1.

OS Streptomyces halstedii.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1944;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-33.

RC STRAIN=JW8 / CECT3310;  
RX MEDLINE=93015685; PubMed=1400190;  
RA Fernandez-Abalos J.M., Sanchez P., Coll-Fresno P.M.,  
RA Villanueva J.R., Perez P., Santamaria R.I.;  
RT "Cloning and nucleotide sequence of celal, and  
endo-beta-1,4-glucanase-encoding gene from Streptomyces halstedii  
JW8.";  
RL J. Bacteriol. 174:6368-6376(1992).

CC -!- FUNCTION: IMPLICATED IN THE MECHANISM OF INDUCTION EXERTED BY  
CC CELLULOSE.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic  
CC linkages in cellulose, lichenin and cereal beta-D-glucans.

CC -!- SIMILARITY: Belongs to cellulase family B (family 6 of glycosyl  
CC hydrolases).

CC -!- CAUTION: It is uncertain whether Met-1 or Met-9 is the initiator.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; Z12157; CAA78145.1; -  
CC HSSP; P26222; 1TML.

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M74162; AAA233909.1; -.
CC DR EMBL; U18997; AAA58016.1; -.
CC DR EMBL; AE000401; AAC76246.1; -.
CC DR PIR; S25281; S25281.
CC DR EcoGene; EGI1514; gltF.
CC KW Signal; Transmembrane; Complete proteome.
CC SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 254 PROTEIN GLTF.
CC FT TRANSMEM 213 229 POTENTIAL.
CC SQ SEQUENCE 254 AA; 26351 MW; 7424D7D1339A5F27 CRC64;

Query Match 67.9%; Score 36; DB 1; Length 254;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
||| |||
DB 45 TPVTGGG 52

RESULT 6
ID ISPE XANAC STANDARD; PRT; 295 AA.
AC Q8PNU1.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)
DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).
GN ISPE OR IPK OR XAC0948.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=20202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Catalyzes the phosphorylation of the position 2 hydroxy
CC group of 4-diphosphocytidyl-2C-methyl-D-erythritol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-
CC erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-
CC methyl-D-erythritol.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; fourth
CC step.
CC -!- SIMILARITY: Belongs to the ispe family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE011725; AAM35836.1; -.
CC DR HAMAP; MF_00061; -.
CC DR InterPro; IPR006204; GHMP_kinase.
CC DR InterPro; IPR004424; ispe.
CC DR Pfam; PF00288; GHMP_kinases; 1.
CC DR TIGRFAMs; TIGR00154; ispe; 1.
CC KW Transferase; Kinase; Isoprene biosynthesis; ATP-binding;
CC Complete proteome.
CC FT NP_BIND 106 116 ATP (POTENTIAL).
CC SQ SEQUENCE 295 AA; 31057 MW; 9123D196F55A873 CRC64;

Query Match 67.9%; Score 36; DB 1; Length 295;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
||| |||
DB 242 TPLRTGGGS 250

RESULT 7
ID ISPE XANCP STANDARD; PRT; 295 AA.
AC Q8PC64.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)
DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).
GN ISPE OR IPK OR XCC0871.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=20202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Catalyzes the phosphorylation of the position 2 hydroxy
CC group of 4-diphosphocytidyl-2C-methyl-D-erythritol (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-
CC erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-
CC methyl-D-erythritol.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; fourth
CC step.
CC -!- SIMILARITY: Belongs to the ispe family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC use by non-profit institutions as long as its content is in no way

```

FT DISULFID 37 154 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 70 86 BY SIMILARITY.  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 223 BY SIMILARITY.  
 FT DISULFID 234 262 BY SIMILARITY.  
 FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 342 AA; 36729 MW; 0620DE88ED187D0F CRC64;  
 Query Match 73.6%; Score 39; DB 1; Length 342;  
 Best Local Similarity 75.0%; Pred. No. 6.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 Db 43 PRITGGGS 50  
 RESULT 4  
 PSS8 RAT STANDARD; PRT; 342 AA.  
 AC Q9ES87; Q9ER01;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prostatin precursor (EC 3.4.21.-).  
 GN PRS8.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Adachi M., Kitamura K., Miyoshi T., Tomita K.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Wang C.;  
 RT "Molecular cloning and expression of rat prostatin.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Possesses a trypsin-like cleavage specificity (By similarity).  
 CC -1- SUBUNIT: Heterodimer of two chains, light and heavy, held by a disulfide bond (By similarity).  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF ITS C-TERMINUS (BY SIMILARITY).  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB017638; BAB20281.1; -  
 CC EMBL; AF202076; AAG32641.1; -  
 CC HSSP; P00734; 1UVS.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR001254; Peptidase\_S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00020; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;  
 KW Transmembrane. 1 29  
 FT SIGNAL 30 32 POTENTIAL.  
 FT PROPEP 33 34 ACTIVATION PEPTIDE (BY SIMILARITY).  
 FT CHAIN 34 44 PROSTATIN LIGHT CHAIN.  
 FT CHAIN 45 45 PROSTATIN HEAVY CHAIN.  
 FT PROPEP 323 342 BY SIMILARITY.  
 FT TRANSMEM 320 340 POTENTIAL.  
 FT DOMAIN 45 286 SERINE PROTEASE.  
 FT DISULFID 37 154 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 70 86 BY SIMILARITY.  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 262 BY SIMILARITY.  
 FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 17 17 I -> V (IN REF. 1).  
 FT CONFLICT 292 292 A -> V (IN REF. 1).  
 SQ SEQUENCE 342 AA; 36843 MW; SED1AF05D9213B98 CRC64;  
 Query Match 73.6%; Score 39; DB 1; Length 342;  
 Best Local Similarity 75.0%; Pred. No. 6.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 Db 43 PRITGGGS 50  
 RESULT 5  
 GLTF\_ECOLI STANDARD; PRT; 254 AA.  
 ID GLTF\_ECOLI  
 AC P28721;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein gltf precursor.  
 GN GLTF OR B3214.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=93078627; PubMed=1447980;  
 RA Castano I., Flores N., Valle F., Covarrubias A.A., Bolivar F.;  
 RT "gltF, a member of the gltBDF operon of Escherichia coli, is involved in nitrogen-regulated gene expression.";  
 RL Mol. Microbiol. 6:2733-2741 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.; Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1234-1244 (1997).  
 CC -1- FUNCTION: INVOLVED IN INDUCTION OF THE SO-CALLED NTR ENZYMES IN RESPONSE TO NITROGEN DEPRIVATION, AS WELL AS IN GLUTAMATE BIOSYNTHESIS. MAY MEDIATE THE GLUTAMATE-DEPENDENT REPRESSION OF THE GLT OPERON.  
 CC -1- SIMILARITY: TO E.COLI YHCF.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB017638; BAB20281.1; -  
 CC EMBL; AF202076; AAG32641.1; -  
 CC HSSP; P00734; 1UVS.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR001254; Peptidase\_S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00020; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;  
 KW Transmembrane. 1 29  
 FT SIGNAL 30 32 POTENTIAL.  
 FT PROPEP 33 34 ACTIVATION PEPTIDE (BY SIMILARITY).  
 FT CHAIN 34 44 PROSTATIN LIGHT CHAIN.  
 FT CHAIN 45 45 PROSTATIN HEAVY CHAIN.  
 FT PROPEP 323 342 BY SIMILARITY.  
 FT TRANSMEM 320 340 POTENTIAL.  
 FT DOMAIN 45 286 SERINE PROTEASE.  
 FT DISULFID 37 154 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 70 86 BY SIMILARITY.  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 262 BY SIMILARITY.  
 FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 17 17 I -> V (IN REF. 1).  
 FT CONFLICT 292 292 A -> V (IN REF. 1).  
 SQ SEQUENCE 342 AA; 36843 MW; SED1AF05D9213B98 CRC64;  
 Query Match 73.6%; Score 39; DB 1; Length 342;  
 Best Local Similarity 75.0%; Pred. No. 6.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 Db 43 PRITGGGS 50

```

SQ  SEQUENCE  561 AA;  62898 MW;  37422EA149E88F30 CRC64;

Query Match      100.0%;  Score 53;  DB 1;  Length 561;
Best Local Similarity 100.0%;  Pred. No. 0.027;
Matches 10;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  1  TPRVTGGGAM 10
Db  417  TPRVTGGGAM 426

RESULT 3
PSS8_MOUSE
ID  PSS8_MOUSE  STANDARD;  PRT;  342 AA.
AC  QRESL;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Proctasin precursor (BC 3.4.21.-) (Channel activating protease 1).
GN  PSS8 OR CAP1.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20235202; PubMed=10770960;
RA  Vaguiaux G., Vallet V., Jaeger N.F., Pfister C., Bens M., Farman N.,
RA  Courtois-Coudry N., Vandewalle A., Rossier B.C., Hummler E.;
RT  "Activation of the amiloride-sensitive epithelial sodium channel by
RT  the serine protease MCAPI expressed in a mouse cortical collecting
RT  duct cell line.";
RL  J. Am. Soc. Nephrol. 11:828-834(2000).
CC  -!- FUNCTION: Possesses a trypsin-like cleavage specificity (By
CC  similarity). Activates amiloride-sensitive sodium channels.
CC  -!- SUBUNIT: Heterodimer of two chains, light and heavy, held by a
CC  disulfide bond (By similarity).
CC  -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF
CC  ITS C-TERMINUS (BY SIMILARITY).
CC  -!- SIMILARITY: Belongs to peptidase family S1.
CC  -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC  frameshift in position 339.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF188613; AAG17054.1; ALT_FRAME.
DR  HSPG; P00734; 1UVS.
DR  MEROPS; S01.158; -.
DR  MGP; MGI:1923810; Prss8.
DR  InterPro; IPR009003; Cys_Ser_trypsin.
DR  InterPro; IPR001254; Peptidase_S1.
DR  InterPro; IPR001314; Peptidase_S1A.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  SMART; SM00020; Tryp_SPC; 1.
DR  PROSITE; PS02040; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;
KW  Transmembrane.
FT  SIGNAL 1 29
FT  PROPEP 30 32
FT  CHAIN 33 44
FT  CHAIN 45 322
FT  PROPEP 323 342
FT  TRANSMEM 320 340
FT  DOMAIN 45 286
FT  SIGNAL 1 29
FT  PROPEP 30 32
FT  CHAIN 33 44
FT  CHAIN 45 322
FT  PROPEP 323 342
FT  TRANSMEM 320 340
FT  DOMAIN 45 286

```

399	28	52.8	158	1	GRP2 ARATH	Q9evm8 arabidopsis	472	28	52.8	354	1	KLIF_MOUSE	O60843 mus musculus
400	28	52.8	162	1	AROK_HELPY	P56073 helicobacte	473	28	52.8	354	1	RNFD_PSEST	Q9evn4 pseudomonas
401	28	52.8	166	1	ING_TURTR	Q9tv67 turlops tr	474	28	52.8	359	1	K6P1_ANASP	O8yk93 anabaena sp
402	28	52.8	170	1	AROK_NEIMA	Q9jqv1 neiseria m	475	28	52.8	367	1	YMP4_STRCO	P43167 streptomyce
403	28	52.8	172	1	AROK_PSEAE	P34003 pseudomonas	476	28	52.8	375	1	HXA2_CHICK	Q08727 gallus gall
404	28	52.8	172	1	AROK_PSEPK	Q88cv1 pseudomonas	477	28	52.8	375	1	PLVB_ASPNG	Q00205 aspergillus
405	28	52.8	172	1	AROK_PSEPM	Q87v14 pseudomonas	478	28	52.8	378	1	YIBH_ECOLI	P32107 escherichia
406	28	52.8	174	1	AROL_YERPE	O8sc15 versinia pe	479	28	52.8	378	1	YIBH_ECOLI	P27693 bacillus al
407	28	52.8	175	1	AROK_PASMU	P57925 pasteurella	480	28	52.8	380	1	ELYA_BACAO	P41362 bacillus cl
408	28	52.8	176	1	YL40_MYCTU	O62235 mycobacteri	481	28	52.8	380	1	ELYA_BACCS	P00374 colletoctric
409	28	52.8	180	1	AROK_HAEIN	P43880 haemophilus	482	28	52.8	380	1	PLVA_COLGL	Q03835 bacillus su
410	28	52.8	200	1	CNTF_PIG	O07732 sus scrofa	483	28	52.8	381	1	SUBN_BACSA	P00783 bacillus su
411	28	52.8	207	1	TRPF_STAEP	O8csn5 staphylococ	484	28	52.8	381	1	SUBT_BACST	P29142 bacillus st
412	28	52.8	208	1	RS5_AOUAE	O67563 aquifex aeo	485	28	52.8	381	1	SUBT_BACST	P04189 bacillus su
413	28	52.8	209	1	COEC_SCYCA	P81130 scyllorhinu	486	28	52.8	386	1	KPBG_HUMAN	P00518 oryctolagus
414	28	52.8	211	1	GSPI_YANCP	P31740 synchocyst	487	28	52.8	386	1	KPBG_HUMAN	Q46889 escherichia
415	28	52.8	214	1	MSA2_SYNY3	P72800 synchocyst	488	28	52.8	388	1	YGBK_ECOLI	Q59002 methanococc
416	28	52.8	217	1	HSB7_MOUSE	P59024 mus musculus	489	28	52.8	390	1	YG07_METJA	Q24350 drosophila
417	28	52.8	219	1	WS18_MOUSE	P53041 mus musculus	490	28	52.8	390	1	YG07_METJA	P27782 mus musculus
418	28	52.8	223	1	PGC2_HUMAN	O15173 homo sapien	491	28	52.8	397	1	LEF1_MOUSE	Q9qnl1 rattus norv
419	28	52.8	227	1	PVBH_SULSO	Q97ze2 sulfolobus	492	28	52.8	397	1	LEF1_MOUSE	Q9juj2 homo sapien
420	28	52.8	228	1	CLDA_HUMAN	P78369 homo sapien	493	28	52.8	399	1	LEF1_MOUSE	Q8rip5 mus musculus
421	28	52.8	233	1	PDUB_SALTY	P37449 salmonella	494	28	52.8	405	1	CIWD_MOUSE	Q9wv69 mus musculus
422	28	52.8	238	1	YODX_BACSU	P14802 bacillus su	495	28	52.8	405	1	CIWD_MOUSE	Q9wv69 mus musculus
423	28	52.8	238	1	AMDA_BORBR	O05115 bordetella	496	28	52.8	405	1	CIWD_MOUSE	Q9wv69 mus musculus
424	28	52.8	240	1	LINC_PSEPA	P50197 pseudomonas	497	28	52.8	405	1	CIWD_MOUSE	Q9wv69 mus musculus
425	28	52.8	253	1	YBPH_APLCA	P1824 aplysia cal	498	28	52.8	405	1	CIWD_MOUSE	Q9wv69 mus musculus
426	28	52.8	263	1	THIM_STAAM	Q99ag5 staphylococ	499	28	52.8	429	1	DADA_XANAC	Q9p9c9 xanthomonas
427	28	52.8	263	1	YF66_ARCFU	O28706 archaeoglob	500	28	52.8	429	1	DADA_XANAC	Q9p4q9 xanthomonas
428	28	52.8	266	1	NRTD_SYNY3	P73265 synchocyst							Q92a14 rhizobium m
429	28	52.8	266	1	THIM_STAAM	Q8nv44 staphylococ							
430	28	52.8	268	1	CH14_BRANA	Q06209 brassica na							
431	28	52.8	268	1	YGLD_ECOLI	Q46927 escherichia							
432	28	52.8	269	1	EL2_PIG	P08419 sus scrofa							
433	28	52.8	269	1	IOD3_RANCA	P49898 rana catesb							
434	28	52.8	269	1	SUBB_BACLE	P29599 bacillus le							
435	28	52.8	269	1	SUBS_BACLE	P29600 bacillus le							
436	28	52.8	271	1	HXD9_ORYLA	Q9pvr2 oryzias lat							
437	28	52.8	275	1	SUBT_BACPU	P07518 bacillus pu							
438	28	52.8	276	1	BPHB_PRES1	P50206 p cis-2,3-d							
439	28	52.8	288	1	HSLO_LACLA	Q9ce86 lactococcus							
440	28	52.8	292	1	AROK_ARATH	Q9sj05 arabidopsis							
441	28	52.8	293	1	RPPI_YEAST	P38786 saccharomyc							
442	28	52.8	297	1	PRTA_STRGR	P00776 streptomyce							
443	28	52.8	297	1	AROK_LYCES	Q00497 lycopersico							
444	28	52.8	300	1	CPPM_BACHD	Q929c7 bacillus ha							
445	28	52.8	303	1	PCPR_SPHCR	P52679 sphingobium							
446	28	52.8	310	1	YN00_MYCTU	Q50665 mycobacteri							
447	28	52.8	311	1	PYRD_BACSU	P25996 bacillus su							
448	28	52.8	317	1	BSS4_HUMAN	Q9gzr4 homo sapien							
449	28	52.8	318	1	CH11_ORYSA	P24626 oryza sativ							
450	28	52.8	325	1	FIBP_ADE1A	P35773 human adeno							
451	28	52.8	325	1	FIBP_ADE1P	P35774 human adeno							
452	28	52.8	326	1	CT77_MOUSE	Q9nqg5 homo sapien							
453	28	52.8	326	1	LDH_TOXGO	Q9csu0 mus musculus							
454	28	52.8	326	1	CEBB_CHICK	Q27797 toxoplasma							
455	28	52.8	328	1	YB94_METH	Q05826 gallus gall							
456	28	52.8	331	1	B3G1_HUMAN	O27262 methanobact							
457	28	52.8	334	1	B3G1_MOUSE	Q9p2w7 homo sapien							
458	28	52.8	334	1	CBFC_MOUSE	O35789 r galactosy							
459	28	52.8	335	1	CBFC_MOUSE	P70353 mus musculus							
460	28	52.8	335	1	CBFC_MOUSE	Q62725 rattus norv							
461	28	52.8	336	1	LPLA_MYGE	P47512 mycoplasma							
462	28	52.8	336	1	RT09_CANAL	O94150 candida alb							
463	28	52.8	338	1	YB61_METH	O27329 methanobact							
464	28	52.8	339	1	LPLA_MYCPN	P75394 mycoplasma							
465	28	52.8	339	1	PCB3_MOUSE	P57722 mus musculus							
466	28	52.8	342	1	RX_MOUSE	O35602 mus musculus							
467	28	52.8	343	1	FIBP_ADE07	P15141 human adeno							
468	28	52.8	343	1	SIAM_MOUSE	Q9gm4 mus musculus							
469	28	52.8	344	1	ARGC_THERH	P96136 thermus the							
470	28	52.8	346	1	YMLC_BACSU	P39153 bacillus su							
471	28	52.8	351	1	KLIF2_RAT	Q9et58 rattus norv							

## ALIGNMENTS

## RESULT 1

PP65\_HCMVT  
ID PP65\_HCMVT STANDARD; PRT; 551 AA.  
AC P18139;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE 64 kDa lower matrix phosphoprotein (PP64) (GP64).  
UL83.  
Human cytomegalovirus (strain Towne).  
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
NCBI\_TaxID=10363;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91220654; PubMed=1850902;  
Pande H., Campo K., Tanamachi B., Zaia J.A.;  
"Human cytomegalovirus strain Towne pp65 gene: nucleotide sequence and expression in Escherichia coli.";  
Virology 182:220-228(1991).  
[2]  
PARTIAL SEQUENCE, AND PHOSPHORYLATION.  
MEDLINE=90357792; PubMed=2167561;  
Pande H., Lee T.D., Churchill M.A., Zaia J.A.;  
"Structural analysis of a 64-kDa major structural protein of human cytomegalovirus (Towne): identification of a phosphorylation site and comparison to pp65 of HCMV (AD169).";  
Virology 178:6-14(1990).  
CC -!- FUNCTION: FORMS PART OF THE MATRIX OF THE HCMV VIRION.  
CC -!- SIMILARITY: BELONGS TO THE UL82 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
-----



253	29	54.7	384	1	GALLI ACTPL	P94169 actinobacil	326	29	54.7	785	1	Y043 METJA	O60348 methanococ
254	29	54.7	384	1	GALLI HAEN	P31767 haemophilus	327	29	54.7	797	1	RBF_DROME	Q24472 drosophila
255	29	54.7	385	1	GALLI PASMU	P57899 pasteurella	328	29	54.7	825	1	ICPO HSV2H	P28284 herpes simp
256	29	54.7	385	1	GUNF_FUSOX	P46239 fusarium ox	329	29	54.7	878	1	CLPB TREPA	O83110 treponema p
257	29	54.7	386	1	GALLI VIBCH	Q9krpl vibrio chol	330	29	54.7	880	1	SYA_IACPL	Q88v10 lactobacill
258	29	54.7	386	1	GALLI VIBPA	Q87m60 vibrio para	331	29	54.7	921	1	ITH4_PIG	P79263 sus scrofa
259	29	54.7	386	1	GALLI VIBVU	Q8dbn9 vibrio vuln	332	29	54.7	930	1	PMP8_CHLPN	Q92393 chlamydia p
260	29	54.7	387	1	GALLI STROCO	Q9k388 streptomyce	333	29	54.7	939	1	SLAP_CAMFE	P35827 campylobact
261	29	54.7	389	1	DUS_MYCTU	O53835 mycobacteri	334	29	54.7	948	1	HP11_DEIRA	P56867 deinococcus
262	29	54.7	391	1	GALLI MOUSE	Q9r0n0 mus musculu	335	29	54.7	1019	1	ENTK HUMAN	P98073 homo sapien
263	29	54.7	392	1	GALLI CANPA	Q95kk4 canis famil	336	29	54.7	1019	1	POL_SIVS4	P12502 simian immu
264	29	54.7	392	1	GALLI HUMAN	P51570 homo sapien	337	29	54.7	1022	1	POL_SIVS4	P19505 simian immu
265	29	54.7	392	1	PRTD STRGR	P52321 streptomyce	338	29	54.7	1035	1	TAC2_MOUSE	Q9ij90 mus musculu
266	29	54.7	394	1	SPEB SCHPO	Q10088 schizosacch	339	29	54.7	1039	1	MSL1_DROME	P50535 drosophila
267	29	54.7	395	1	HLK_VIBCH	Q9kv09 vibrio chol	340	29	54.7	1054	1	POL_SIVM1	P50535 drosophila
268	29	54.7	395	1	RAPP_NDVA	P16073 newcastle d	341	29	54.7	1056	1	POL_SIVM1	P50535 drosophila
269	29	54.7	395	1	RAPP_NDVA	P24698 newcastle d	342	29	54.7	1058	1	POL_HV2D2	P50535 drosophila
270	29	54.7	395	1	RAPP_NDVU2	Q06427 newcastle d	343	29	54.7	1069	1	S24E_ARATH	P50535 drosophila
271	29	54.7	395	1	FLO_ANTMA	P23915 antirrhinum	344	29	54.7	1073	1	POL_HV2D1	P17757 human immu
272	29	54.7	396	1	GALLI STRLI	P13227 streptomyce	345	29	54.7	1088	1	GP22_DROME	Q03043 drosophila
273	29	54.7	397	1	DROB_CLOTE	Q89422 clostridium	346	29	54.7	1097	1	S24C_ARATH	Q9m291 arabidopsis
274	29	54.7	398	1	FRSZ_WOLSP	P45485 wolbachia s	347	29	54.7	1098	1	TSL1_YEAST	P38427 saccharomyc
275	29	54.7	398	1	VCAB_BPT7	P19727 bacterioph	348	29	54.7	1142	1	POL_HV2BE	Q08236 saccharomyc
276	29	54.7	400	1	CSD_CHLMU	Q9p1p0 chlamydia m	349	29	54.7	1176	1	YOH8_YEAST	P10745 homo sapien
277	29	54.7	401	1	CSD_CHLTR	Q01889 trypanosoma	350	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
278	29	54.7	405	1	PAG1_TRYBB	Q94715 chlamydia p	351	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
279	29	54.7	406	1	TRYQ_HAEN	P44747 haemophilus	352	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
280	29	54.7	406	1	TRYQ_HAEN	P50317 sulfolobus	353	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
281	29	54.7	408	1	PGK_SULSO	Q9hj95 thermoplasma	354	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
282	29	54.7	408	1	PGK_THEAC	Q9hbc6 thermoplasma	355	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
283	29	54.7	408	1	PGK_THEVO	Q9yfs7 aeropyrum p	356	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
284	29	54.7	410	1	PGK_AERPE	Q9yfs7 aeropyrum p	357	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
285	29	54.7	415	1	PGK_SULTO	Q9yfs7 aeropyrum p	358	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
286	29	54.7	422	1	NCAP_CVHSA	P59595 human coron	359	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
287	29	54.7	425	1	CCGB_HUMAN	Q9qy17 mus musculu	360	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
288	29	54.7	432	1	PT22_MOUSE	P19728 bacterioph	361	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
289	29	54.7	433	1	VCAB_BPT3	P28088 bos taurus	362	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
290	29	54.7	441	1	ETBR_BOVIN	Q8d2x5 streptococ	363	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
291	29	54.7	444	1	GID_STRA3	Q8d2x5 streptococ	364	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
292	29	54.7	444	1	GID_STRA5	Q8d2x5 streptococ	365	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
293	29	54.7	444	1	GID_STRMU	P05428 streptococ	366	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
294	29	54.7	448	1	GID_STRP8	Q8k7g4 streptococ	367	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
295	29	54.7	448	1	GID_STRP8	Q8p106 streptococ	368	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
296	29	54.7	448	1	GID_STRPY	Q99219 streptococ	369	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
297	29	54.7	451	1	CP51_MYCTU	P77901 mycobacteri	370	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
298	29	54.7	451	1	GASH_MOUSE	P56481 mus musculu	371	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
299	29	54.7	453	1	VG_DROME	Q26366 drosophila	372	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
300	29	54.7	468	1	BMP3_RAT	P49002 rattus norv	373	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
301	29	54.7	471	1	V51K_BMYVG	P09515 beet wester	374	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
302	29	54.7	474	1	TOCC_MAIZE	Q94fy8 zea mays (m	375	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
303	29	54.7	476	1	CHK1_MOUSE	O35280 mus musculu	376	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
304	29	54.7	488	1	TOCC_ARATH	Q94fy7 arabidopsis	377	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
305	29	54.7	491	1	HHIR_BOVIN	P30546 bos taurus	378	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
306	29	54.7	498	1	GSHR_PEA	Q4321 pismus sativ	379	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
307	29	54.7	503	1	SIZ_BPSPP	P54309 bacterioph	380	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
308	29	54.7	508	1	GLPK_STRP3	Q8k665 streptococ	381	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
309	29	54.7	508	1	GLPK_STRP8	Q8nzw9 streptococ	382	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
310	29	54.7	508	1	GLPK_STRPY	Q99v17 streptococ	383	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
311	29	54.7	541	1	PSN_DROME	Q02194 drosophila	384	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
312	29	54.7	544	1	CH63_RHIME	Q930y0 rhizobium m	385	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
313	29	54.7	553	1	C301_DROME	Q9v6d6 drosophila	386	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
314	29	54.7	557	1	PAC1_PSES3	P15557 pseudomonas	387	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
315	29	54.7	557	1	PAC1_PSES3	Q05053 pseudomonas	388	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
316	29	54.7	562	1	NUPL_HUMAN	P52594 homo sapien	389	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
317	29	54.7	562	1	HLV4_AERSA	Q08677 aeromonas s	390	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
318	29	54.7	597	1	IXRI_YEAST	P33417 saccharomyc	391	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
319	29	54.7	606	1	GAB_DROME	P25123 drosophila	392	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
320	29	54.7	607	1	RPOL_BPM2	O64235 mycobacteri	393	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
321	29	54.7	607	1	RPOL_HALN1	P15351 halobacteri	394	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
322	29	54.7	618	1	CHAA_CLOBI	O05102 clostridium	395	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
323	29	54.7	625	1	FAIL_HUMAN	P03951 homo sapien	396	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
324	29	54.7	629	1	ODP2_ECOLI	P06959 escherichia	397	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila
325	29	54.7	657	1	Y023_NPVP	O10282 orgyia pseu	398	29	54.7	1247	1	IRBP_HUMAN	P13055 drosophila





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:01:12 ; Search time 11 Seconds  
(without alignments)

47.337 Million cell updates/sec

Title: US-10-697-055-7

Perfect score: 53

Sequence: 1 TPRVTGGGAM 10

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	53	100.0	551	1	PP65 HCMVT
2	53	100.0	561	1	PP65 HCMVA
3	39	73.6	342	1	PSS8 MOUSE
4	39	73.6	342	1	PSS8 RAT
5	36	67.9	254	1	GLTP ECOLI
6	36	67.9	295	1	ISPE XANAC
7	36	67.9	295	1	ISPE XANCP
8	36	67.9	799	1	GLGB MALZE
9	35	66.0	321	1	GUNI STRHA
10	35	66.0	345	1	GAS1 HUMAN
11	35	66.0	359	1	CRL2 MOUSE
12	35	66.0	524	1	PIF3 ARATH
13	34	64.2	206	1	MUSC HUMAN
14	34	64.2	247	1	YGGH YEAST
15	34	64.2	269	1	THIM LISIN
16	34	64.2	269	1	THIM LISMO
17	34	64.2	380	1	COAT CNV
18	34	64.2	408	1	YS92 MYCTU
19	33	62.3	223	1	VZL2 EBV
20	33	62.3	224	1	RS5 YEAST
21	33	62.3	260	1	PSAI CAEBL
22	33	62.3	311	1	SDCI MOUSE
23	33	62.3	345	1	RL4 CAEBL
24	33	62.3	386	1	RL4 URECA
25	33	62.3	396	1	RL4A XENLA
26	33	62.3	396	1	RL4B XENLA
27	33	62.3	406	1	RL4B ARATH
28	33	62.3	407	1	RL4A ARATH
29	33	62.3	407	1	RL4 DROME
30	33	62.3	408	1	RL4 PRUAR
31	33	62.3	408	1	SEPR THESR
32	33	62.3	419	1	RL4 MOUSE
33	33	62.3	420	1	RL4 CANFA

34	62.3	421	1	RL4 RAT	P50878 rattus norv
35	62.3	427	1	RL4 HUMAN	P36578 homo sapien
36	62.3	435	1	TMS4 MOUSE	Q8vca5 mus musculu
37	62.3	437	1	TMS4 HUMAN	Q9nrs4 homo sapien
38	62.3	444	1	SHU4 ECOLI	P09748 escherichia
39	62.3	476	1	CHK1 HUMAN	O14757 homo sapien
40	62.3	487	1	OAF DROME	Q9nla6 drosophila
41	62.3	561	1	AIAD RAT	P23944 rattus norv
42	62.3	663	1	PD11 HUMAN	Q9ulc6 homo sapien
43	62.3	692	1	EFG STRAP	P54741 streptomyce
44	62.3	799	1	AFSK STRCO	P54742 streptomyce
45	62.3	807	1	AFSK STRGR	P54742 streptomyce
46	62.3	870	1	PNK1 ARATH	O815Y9 arabidopsis
47	62.3	1036	1	HPI2 DEIRA	P13126 deinococcus
48	62.3	1093	1	AF17 HUMAN	P55198 homo sapien
49	62.3	1244	1	SLA1 YEAST	P32790 saccharomyc
50	62.3	142	1	RS16 SACEX	Q876b4 saccharomyc
51	62.3	142	1	RS16 YEAST	P40213 saccharomyc
52	62.3	196	1	MIP TRYCR	Q08734 trypanosoma
53	62.3	206	1	SMGH XANAC	Q8pj11 xanthomonas
54	62.3	230	1	YGTI ECOLI	Q46844 escherichia
55	62.3	262	1	LIPB THEAC	Q9nht1 thermoplasm
56	62.3	264	1	LIPB AERPE	Q9v9e6 aeropyrum p
57	62.3	271	1	CTR1 PENVA	Q00871 penaeus van
58	62.3	271	1	CTR2 PENVA	P36178 penaeus van
59	62.3	317	1	MOCA RHIME	P49307 rhizobium m
60	62.3	333	1	MEL BACSU	P39751 bacillus su
61	62.3	336	1	UL34 EBV	P03185 epstein-bar
62	62.3	361	1	RL4A YEAST	P10664 saccharomyc
63	62.3	361	1	RL4B YEAST	P49626 saccharomyc
64	62.3	363	1	RL4A SCHPO	P35679 schizosacch
65	62.3	363	1	RL4B SCHPO	Q9P784 schizosacch
66	62.3	406	1	PGK METKA	Q8tuul methanopyru
67	62.3	410	1	PGK PYRAB	Q9uzw0 pyrococcus
68	62.3	410	1	PGK PYRFU	P50316 pyrococcus
69	62.3	410	1	PGK PYRHO	Q58965 pyrococcus
70	62.3	419	1	NIFS FRASE	Q925x5 frankia sp.
71	62.3	432	1	APRE PSEAE	Q03025 pseudomonas
72	62.3	434	1	YAOA SCHPO	Q10089 schizosacch
73	62.3	467	1	CLCA VIBPA	Q87g29 vibrio para
74	62.3	472	1	EG44 CAEBL	Q19489 caenorhabdi
75	62.3	491	1	NMT CRANE	P34809 cryptococcu
76	62.3	604	1	FUCI HABIN	P44779 haemophilus
77	62.3	687	1	FSHR EQUAS	Q95179 equus asinu
78	62.3	827	1	PTF1 ROCA	P23388 r multiphos
79	62.3	838	1	L100 ADEP3	Q9vyr7 porcine ade
80	62.3	1180	1	RNT1 DROME	Q9vyr3 drosophila
81	62.3	91	1	NLT1 PRUPE	P81402 prunus pers
82	62.3	117	1	NLT1 PRUDU	Q43017 prunus dulc
83	62.3	151	1	RS9 AERPE	Q9y448 aeropyrum p
84	62.3	170	1	Y5A3 VIBPA	Q87125 vibrio para
85	62.3	215	1	RL4 MYCSM	Q06114 mycobacteri
86	62.3	223	1	RL4 MYCTU	P95050 mycobacteri
87	62.3	275	1	BNZE PSEPU	P08088 pseudomonas
88	62.3	275	1	TODD PSEPU	P13859 pseudomonas
89	62.3	299	1	CC34 CAEBL	P34867 caenorhabdi
90	62.3	312	1	EBP2 XENLA	Q918j6 xenopus lae
91	62.3	325	1	UL76 HCMVA	P16725 human cytom
92	62.3	346	1	NRL3 ARATH	P46010 arabidopsis
93	62.3	368	1	PPBL PSEAE	P35482 pseudomonas
94	62.3	369	1	DNJ1 THEMA	Q9wz3 thermotoga
95	62.3	370	1	PST3 MYCTU	Q86343 mycobacteri
96	62.3	374	1	RL4 TRYBB	P49669 trypanosoma
97	62.3	380	1	POLG MDV	P32652 maize dwarf
98	62.3	410	1	ISCS RICCN	Q92hpl rickettsia
99	62.3	410	1	ISCS RICPR	Q92d60 rickettsia
100	62.3	410	1	PEPT LISIN	Q92am8 listeria in
101	62.3	410	1	PEPT LISMO	Q8y6b1 listeria mo
102	62.3	426	1	GLYA THEAC	Q9h138 thermoplasm
103	62.3	450	1	YHXA BACSU	P33189 bacillus su
104	62.3	451	1	RADA TREPA	Q83985 treponema p
105	62.3	458	1	GNFR HUMAN	Q9ugk8 homo sapien
106	62.3	467	1	CLCA VIBVU	Q8d6j0 vibrio vuln

Best Local Similarity 77.8%; Pred. No. 1.4e+03;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPVRTGGGA 9  
DB 778 TPRATGRGA 786

PRELIMINARY; PRT; 3456 AA.

RESULT 49  
P89201  
ID P89201  
AC P89201  
DT 01-MAY-1997 (Tremblrel. 03, Created)  
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Polyprotein.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
SEQUENCE FROM N.A.  
MEDLINE=98281520; PubMed=9620210;  
Colinet D., Kummert J., Lepoivre P.;  
"The nucleotide sequence and genome organization of the whitefly  
transmitted sweetpotato mild mottle virus: a close relationship with  
members of the family Potyviridae.";  
RT Virus Res. 53:187-196(1998).  
RL EMBL; Z73124; CAA97466.1; --  
DR MEROPS; C04.013; --  
DR MEROPS; C06.001; --  
DR GO; GO:0019028; C:Viral capsid; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.  
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0016787; F:hydroxylase activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0019079; P:viral genome replication; IEA.  
DR InterPro; IPR001410; DEAD  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR001730; Peptidase\_C4.  
DR InterPro; IPR001456; Peptidase\_C6.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF00271; helicase\_C; 1.  
DR Pfam; PF00863; Peptidase\_C4; 1.  
DR Pfam; PF00851; Peptidase\_C6; 1.  
DR Pfam; PF00767; Poty coat; 1.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR PRINTS; PR00966; NIAPOTYPASE.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELICC; 1.  
KW ATP-binding; Coat protein; Helicase; Hydrolase.  
FT CHAIN 1 743 POTENTIAL.  
FT CHAIN 744 863 HELPER COMPONENT.  
FT CHAIN 864 1489 POTENTIAL.  
FT CHAIN 1490 1540 6K1.  
FT CHAIN 1541 2181 HELICASE.  
FT CHAIN 2182 3456 NUCLEAR INCLUSION PROTEIN A/B, COAT  
PROTEIN.  
SQ SEQUENCE 3456 AA; 390279 MW; C8D758543597C72A CRC64;

Query Match 67.9%; Score 36; DB 12; Length 3456;  
Best Local Similarity 77.8%; Pred. No. 3.1e+03;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TPVRTGGGA 10

Db 3224 PRVTGFGAL 3232

PRELIMINARY; PRT; 126 AA.

RESULT 50  
Q9CRJ6  
ID Q9CRJ6  
AC Q9CRJ6  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE Thymic stromal-derived lymphopoietin, receptor (Fragment).  
GN TSLPR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;  
MEDLINE=21085660; PubMed=11217851;  
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
Saito I., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
Fleischmann W., Gaasterland T., Gissi C., King B., Quackenbush J.,  
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Wagner L., Washio T.,  
Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Barsh G.,  
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
Northone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
Hayashizaki Y.;  
"Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK010291; BAB26827.1; --  
DR MGD; MGI:1889506; Tslpr.  
DR NON\_TER 1  
FT SEQUENCE 126 AA; 13321 MW; BD387D755A8D9DF6 CRC64;

Query Match 66.0%; Score 35; DB 11; Length 126;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGA 10  
DB 100 PRGPGGGGA 108

Search completed: March 7, 2004, 13:02:29  
Job time : 63 secs

Query Match 67.9%; Score 36; DB 2; Length 808;  
 Best Local Similarity 75.0%; Pred. No. 7.3e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 176 TPGITGGG 183

RESULT 46

Q8GN89

ID Q8GN89

AC Q8GN89

DT 01-JUN-2003

DT 01-WAR-2003

DT 01-WAR-2003

DT 01-JUN-2003

DE FpVA11.

GN FpVA.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI\_TaxID=287;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=7NSK2;

RX MEDLINE=9717770; PubMed=9025276;

RA Meyer J.M., Stintzi A., De Vos D., Cornelis P., Tappe R., Taraz K.,

RA Budzikiewicz H.;

RT "Use of siderophores to type pseudomonads: the three Pseudomonas

RT aeruginosa pyoverdine systems.";

RL Microbiology 143:35-43(1997).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=7NSK2;

RA De Chial M., Ghysels B., Beatson S., Geoffroy V., Meyer J.-M.,

RA Pattery T., Bayasse C., Martens J., Chablain P., Parsons Y.N.,

RA Winstanley C., Cordwell S., Cornelis P.;

RT "Identification of the two undescribed ferripyoverdine receptors from

RT Pseudomonas aeruginosa.";

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF537095; AA62913.1;

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:004872; F:receptor activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR00531; TonB\_boxC.

DR Pfam; PF00593; TonB\_dep\_Rec; I.

SQ SEQUENCE 823 AA; 91077 MW; 7848FD9F10FIDCA2 CRC64;

Query Match 67.9%; Score 36; DB 2; Length 823;

Best Local Similarity 75.0%; Pred. No. 7.4e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 191 TPGITGGG 198

RESULT 47

Q89UL2

ID Q89UL2

AC Q89UL2

DT 01-JUN-2003

DT 01-JUN-2003

DT 01-OCT-2003

DT 01-OCT-2003

DE ATP-dependent protease ATP-binding subunit.

GN CLPB OR BLR1404.

OS Bradyrhizobium japonicum.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Bradyrhizobium.

OX NCBI\_TaxID=375;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=USDA 110;  
 RX MEDLINE=2248498; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
 RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tabata S.;

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium

RT Bradyrhizobium japonicum USDA110.";

RL DNA Res. 9:189-197(2002).

DR EMBL; AP005940; BAC46669.1; --

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003754; F:chaperone activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR InterPro; IPR003959; AAA ATPase centr.

DR InterPro; IPR001270; Chaprinin\_clpA/B.

DR InterPro; IPR004176; Clp\_N.

DR Pfam; PF00004; AAA; 1.

DR Pfam; PF02861; Clp\_N; 2.

DR PRINTS; PR00300; CLPPTPTEASEA.

DR PROSITE; PS00870; CLPAB\_1; 1.

DR PROSITE; PS00871; CLPAB\_2; 1.

KW ATP-binding; Protease; Complete proteome.

SQ SEQUENCE 879 AA; 96621 MW; 0F935D239D7A3867 CRC64;

Query Match 67.9%; Score 36; DB 16; Length 879;

Best Local Similarity 75.0%; Pred. No. 7.9e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9

DB 72 PKVSGGGA 79

RESULT 48

OL5080

ID OL5080

AC OL5080

DT 01-JAN-1998

DT 01-MAY-2000

DT 01-OCT-2003

DE Hypothetical protein KIAA0375 (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97349984; PubMed=9205841;

RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,

RA Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. VII.

RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro.";

RL DNA Res. 4:141-150(1997).

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

DR EMBL; AB002373; BAA20830.2; --

DR HSP; O61259; IABO.

DR InterPro; IPR004012; Run.

DR InterPro; IPR001452; SH3.

DR Pfam; PF02759; RUN; 1.

DR Pfam; PF00018; SH3; 1.

DR ProDom; PD000066; SH3; 1.

DR SMART; SM00593; RUN; 1.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PSS0826; RUN; 1.

DR PROSITE; PSS0002; SH3; 1.

KW Hypothetical protein; SH3 domain.

FT NON TER 1

SQ SEQUENCE 1540 AA; 163853 MW; 5CC4D2C8FE77083B CRC64;

Query Match

67.9%; Score 36; DB 4; Length 1540;

Sun Mar 7 13:59:05 2004

RC STRAIN=ISH;  
RA Mukasa S.B.;  
RT "Analysis of the 3' terminal genomic region of sweet potato mild  
RT mottle virus isolates from Uganda."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AJ459318; CAD30642.1; -;  
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0019079; P:Viral genome replication; IEA.  
DR InterPro; IPR001592; PotY\_coat.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF00767; PotY\_coat; 1.  
FT NON\_TER 1 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.  
SQ SEQUENCE 501 AA; 57263 MW; DDC933F71AD0B91A CRC64;  
Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGM 10  
DB 269 PRVTGGGM 277  
RESULT 43  
O67557  
ID O67557 PRELIMINARY; PRT; 788 AA.  
AC O67557  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Lipoate-protein ligase A.  
GN LPLA OR AQ\_1638.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=VFS;  
RC MEDLINE=98196666; PubMed=9537320;  
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Auja M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus."  
RL Nature 392:353-358(1998).  
DR EMBL; AE000748; AAC07512.1; -;  
DR PIR; C70441; C70441.  
DR GO; GO:0016874; P:ligase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR004143; BPL\_LipA\_LipB.  
DR InterPro; IPR003704; CO\_dh/CoA\_synth.  
DR Pfam; PF03099; BPL\_LipA\_LipB; 1.  
DR Pfam; PF02552; CO\_dh; 1.  
KW Ligase; Complete proteome.  
SQ SEQUENCE 788 AA; 90804 MW; 3303E372FC719320 CRC64;  
Query Match 67.9%; Score 36; DB 16; Length 788;  
Best Local Similarity 75.0%; Pred. No. 7.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 RVTTGGGM 10  
DB 134 RVTTGGGM 141  
RESULT 44  
O81387  
ID O81387 PRELIMINARY; PRT; 799 AA.  
AC O81387  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Starch branching enzyme IIB.  
GN AE.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. B73;  
RA Kim K.-N., Fisher D.K., Gao M., Gultinan M.J.;  
RT "Molecular cloning and characterization of the amylose-extender gene  
RT encoding starch branching enzyme IIB in maize."  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF072725; AAC33764.1; -;  
DR PIR; T01663; T01663.  
DR GO; GO:0004556; F:alpha-amylose activity; IEA.  
DR GO; GO:0005507; F:copper ion binding; IEA.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.  
DR GO; GO:0005975; F:carbohydrate metabolism; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR006047; Alpha\_aml\_cat.  
DR InterPro; IPR000923; BlueCu\_1.  
DR InterPro; IPR004193; GlycoHydro\_13N.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00128; alpha-amylose; 1.  
DR Pfam; PF02922; isoamylose N; 1.  
DR PROSITE; PS00196; COPPER\_BLUP; 1.  
SQ SEQUENCE 799 AA; 90665 MW; 4FD7B1EBC3B759F0 CRC64;  
Query Match 67.9%; Score 36; DB 10; Length 799;  
Best Local Similarity 85.7%; Pred. No. 7.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PRVTGGG 8  
DB 17 PRVTGGG 23  
RESULT 45  
ID Q8G8W9 PRELIMINARY; PRT; 808 AA.  
AC Q8G8W9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Fpva.  
GN Fpva.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1-60, and 2-164;  
RA Spencer D., Kas A., Smith E., Raymond C., Sims E., Hastings M.,  
RA Burns J., Kaul R., Olson M.;  
RT "Whole Genome Sequence Variation Among Multiple Isolates of Pseudomonas  
RT aeruginosa."  
RL J. Bacteriol. 0:0-0(2003).  
DR EMBL; AF540992; AA017428.1; -;  
DR EMBL; AF540993; AA017439.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000531; TonB\_boxC.  
DR Pfam; PF00593; TonB\_dep\_Rec; 1.  
SQ SEQUENCE 808 AA; 89414 MW; 4C7972A4D310E25C CRC64;

SQ SEQUENCE 501 AA; 57173 MW; 89DD08B244DBECID CRC64;  
Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
Db 269 PRVTGGGAL 277  
|||||  
RESULT 39  
Q8B327 PRELIMINARY; PRT; 501 AA.  
ID Q8B327  
AC Q8B327;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Coat protein (Fragment).  
GN CP.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BNY;  
RA Mukasa S.B.;  
RT "Analysis of the 3' terminal genomic region of sweet potato mild mottle virus isolates from Uganda."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ459314; CAD30638.1;  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0019079; P:Viral genome replication; IEA.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF00767; Poty\_coat; 1.  
FT NON\_TER 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.  
SQ SEQUENCE 501 AA; 57182 MW; F06CD1D0061D3DB1 CRC64;  
Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
Db 269 PRVTGGGAL 277  
|||||  
RESULT 40  
Q8B326 PRELIMINARY; PRT; 501 AA.  
ID Q8B326  
AC Q8B326;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Coat protein (Fragment).  
GN CP.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RUK;  
RA Mukasa S.B.;  
RT "Analysis of the 3' terminal genomic region of sweet potato mild mottle virus isolates from Uganda."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ459315; CAD30639.1;  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

DR GO; GO:0019079; P:Viral genome replication; IEA.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF00767; Poty\_coat; 1.  
FT NON\_TER 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.  
SQ SEQUENCE 501 AA; 57371 MW; D2C1967FE1EF5A27 CRC64;  
Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
Db 269 PRVTGGGAL 277  
|||||  
RESULT 41  
Q8B325 PRELIMINARY; PRT; 501 AA.  
ID Q8B325  
AC Q8B325;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Coat protein (Fragment).  
GN CP.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RUK2;  
RA Mukasa S.B.;  
RT "Analysis of the 3' terminal genomic region of sweet potato mild mottle virus isolates from Uganda."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ459316; CAD30640.1;  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0019079; P:Viral genome replication; IEA.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF00767; Poty\_coat; 1.  
FT NON\_TER 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.  
SQ SEQUENCE 501 AA; 57204 MW; FF524F496A865F92 CRC64;  
Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PRVTGGGAM 10  
Db 269 PRVTGGGAL 277  
|||||  
RESULT 42  
Q8B323 PRELIMINARY; PRT; 501 AA.  
ID Q8B323  
AC Q8B323;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Coat protein (Fragment).  
GN CP.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.

DR Gramene; Q9ARP7; -  
SQ SEQUENCE 383 AA; 40697 MW; D0336CA0423584CF CRC64;

Query Match 67.9%; Score 36; DB 10; Length 383;  
Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
| : : : :  
DB 214 TPRTSGGG 221

## RESULT 36

O69817 PRELIMINARY; PRT; 388 AA.  
AC O69817;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein SC06423.  
GN SC06423 OR SC1A6.12C.  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Parkhill J., Bentley S.D., Barrell B.G., Rajandream M.A.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,  
Kinaishi H., Hopwood D.A.;  
RL "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,  
Kinaishi H., Hopwood D.A.;  
RL "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
Thompson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,  
Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,  
Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,  
Hopwood D.A.;  
RL "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2).";  
RN Nature 417:141-147(2002).  
DR EMBL; AL339127; CAAL18910.1; -.  
DR F1R; T28693; T28693.  
DR GO; GO:0003824; P: catalytic activity; IEA.  
DR GO; GO:0006464; P: protein modification; IEA.  
DR InterPro; IPR004143; BPL LipA LipB.  
DR Pfam; PF03099; BPL LipA LipB; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 388 AA; 41814 MW; 9AA1DE930DF6FCA CRC64;

Query Match 67.9%; Score 36; DB 16; Length 388;  
Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RVTGGGAM 10  
| : : : :  
DB 205 RISGGGAM 212

## RESULT 37

Q8B329 PRELIMINARY; PRT; 501 AA.  
ID Q8B329;  
AC Q8B329;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Polypeptide (Fragment).  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TOR;  
RA Mukasa S.B.;  
RL "Analysis of the 3' terminal genomic region of sweet potato mild  
mottle virus isolates from Uganda";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ459312; CAD30636.1; -.  
DR GO; GO:0003968; P: RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0019079; P: viral genome replication; IEA.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007094; RNA\_pol\_Psvir.  
DR Pfam; PF00767; Poty\_coat; 1.  
RN NON\_TER 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.  
SQ SEQUENCE 501 AA; 57249 MW; 77EF5647FB966A6C CRC64;

Query Match 67.9%; Score 36; DB 12; Length 501;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
| : : : :  
DB 269 PRVTGGGAL 277

## RESULT 38

Q8B328 PRELIMINARY; PRT; 501 AA.  
ID Q8B328;  
AC Q8B328;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Coat protein (Fragment).  
GN CP.  
OS Sweet potato mild mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
OC Ipomovirus.  
OX NCBI\_TaxID=41459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KUM;  
RA Mukasa S.B.;  
RL "Analysis of the 3' terminal genomic region of sweet potato mild  
mottle virus isolates from Uganda";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ459313; CAD30637.1; -.  
DR GO; GO:0003968; P: RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0019079; P: viral genome replication; IEA.  
DR InterPro; IPR001592; Poty coat.  
DR InterPro; IPR007094; RNA\_pol\_Psvir.  
DR Pfam; PF00767; Poty\_coat; 1.  
RN NON\_TER 1  
FT CHAIN 227 501 COAT PROTEIN.  
FT CHAIN 1 226 NUCLEAR INCLUSION B.

RX MEDLINE=20117992; PubMed=10652093;  
 RA Loessner M.J., Inman R.B., Lauer P., Calendar R.;  
 RT "Complete nucleotide sequence, molecular analysis and genome structure  
 of bacteriophage A118 of *Listeria monocytogenes*: implications for  
 RT phage evolution.";   
 RL Mol. Microbiol. 35:324-340(2000).  
 DR EMBL; AJ242593; CAB53832.1; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR005039; Anti\_rep.  
 DR Pfam; PF03374; Anti; 1.  
 SQ SEQUENCE 262 AA; 30217 MW; 11B66231BDB6FF440 CRC64;  
  
 Query Match 67.9%; Score 36; DB 9; Length 262;  
 Best Local Similarity 87.5%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 TPRVTGGG 8  
 Db 238 TPRVTGGK 245  
  
 RESULT 33  
 Q9RIP7 PRELIMINARY; PRT; 354 AA.  
 ID Q9RIP7  
 AC Q9RIP7  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE SNO.  
 GN SNO.  
 OS Streptomyces nogalater.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=38314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27451;  
 RX MEDLINE=96242142; PubMed=8668120;  
 RA Ylihonko K., Tuikka J., Jussila S., Cong L., Mantsala P.;  
 RT "A gene cluster involved in nogalater biosynthesis from Streptomyces  
 RT nogalater: sequence analysis and complementation of early-blocked  
 RT mutations in the anthracycline pathway.";   
 RL Mol. Gen. Genet. 251:113-120(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27451;  
 RX MEDLINE=96349102; PubMed=8760909;  
 RA Ylihonko K., Hakala J., Kunnari T., Mantsala P.;  
 RT "Production of hybrid anthracycline antibiotics by heterologous  
 RT expression of Streptomyces nogalater nogalater biosynthesis genes.";   
 RL Microbiology 142:1965-1972(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27451;  
 RX MEDLINE=98007868; PubMed=9349712;  
 RA Torckell S., Ylihonko K., Hakala J., Skurnik M., Mantsala P.;  
 RT "Characterization of Streptomyces nogalater genes encoding enzymes  
 RT involved in glycosylation steps in nogalater biosynthesis.";   
 RL Mol. Gen. Genet. 256:203-209(1997).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27451;  
 RA Palmu K., Hautala A., Kunnari T., Hakala J., Mantsala P., Ylihonko K.;  
 RT "Identification and functional analysis of a glycosyltransferase gene  
 RT (snogZ) and a resistance gene (snorO) from the nogalater producer,  
 RT Streptomyces nogalater.";   
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ224512; CAB59005.1; -;  
 DR PIR; T46683;  
 DR GO; GO:0003824; F:Catalytic activity; IEA.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR004143; BPL\_LipA\_LipB.  
 DR Pfam; PF03099; BPL\_LipA\_LipB; 1.

SQ SEQUENCE 354 AA; 38074 MW; 4C7674ELF23741DD CRC64;  
 Query Match 67.9%; Score 36; DB 2; Length 354;  
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 3 RVTGGGAM 10  
 Db 171 RISGGGAM 178  
  
 RESULT 34  
 Q8G501 PRELIMINARY; PRT; 361 AA.  
 ID Q8G501  
 AC Q8G501  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Probable lipote protein ligase.  
 GN SNO OR BL1217.  
 OS Bifidobacterium longum.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 OC Bifidobacteriaceae; Bifidobacterium.  
 OX NCBI\_TaxID=216816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCC 2705;  
 RX MEDLINE=22294977; PubMed=12381787;  
 RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,  
 RA Pessi G., Zwielen M.-C., Desiere F., Bork P., Delley M.,  
 RA Pridmore R.D., Arigoni F.;  
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
 RT to the human gastrointestinal tract.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
 DR EMBL; AS014748; AAN25024.1; -;  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0006464; P:protein modification; IEA.  
 DR InterPro; IPR004143; BPL\_LipA\_LipB.  
 DR Pfam; PF03099; BPL\_LipA\_LipB; 1.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 361 AA; 39417 MW; 2825C6075B661A38 CRC64;  
  
 Query Match 67.9%; Score 36; DB 16; Length 361;  
 Best Local Similarity 87.5%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 RVTGGGAM 10  
 Db 215 RCTGGGAM 222  
  
 RESULT 35  
 Q9ARP7 PRELIMINARY; PRT; 383 AA.  
 ID Q9ARP7  
 AC Q9ARP7  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE OSJNBa0010K01.19 protein.  
 GN OSJNBa0010K01.19.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone:OSJNBa0010K01.1";   
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003210; BAB40104.1; -;.

OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.,  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531 (2003).  
 DR EMBL; AF005050; BAC75239.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 244 AA; 26824 MW; E953E6C9886999ABD CRC64;  
 Query Match 57.9%; Score 36; DB 16; Length 244;  
 Best Local Similarity 70.0%; Pred. No. 2.2e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 74 TPRVTGGGAM 83  
 RESULT 31  
 Q9W3C6 PRELIMINARY; PRT; 261 AA.  
 ID Q9W3C6  
 AC Q9W3C6  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE CG11294 protein.  
 GN CG11294.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,  
 RA Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delecker A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.-J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; AE003444; AAF46405.1; -.  
 DR HSSP; P06601; 1FJL.  
 DR FlyBase; FBgn0030058; CG11294.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0006491; E:oxidoreductase activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR002086; Aldehyde dehydr.  
 DR InterPro; IPR001356; Homeobox\_  
 DR InterPro; IPR007104; Paired homeo.  
 DR Pfam; PF00046; Homeobox; 1.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00687; ALDEHYDE DEHYDR\_GLU; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00711; HOMEBOX\_2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 261 AA; 28643 MW; 27A9BF152F372FF3 CRC64;  
 Query Match 67.9%; Score 36; DB 5; Length 261;  
 Best Local Similarity 70.0%; Pred. No. 2.4e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 DB 101 TPRVTGGGAM 110  
 RESULT 32  
 Q9T178 PRELIMINARY; PRT; 262 AA.  
 ID Q9T178  
 AC Q9T178  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative anti-repressor promoter.  
 OS Bacteriophage A118.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=40521;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FROM N.A.  
 RX Loessner M.J.;  
 RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96020653; PubMed=8577256;  
 RA Loessner M.J., Wendlinger G., Scherer S.;  
 RT "Heterogeneous endolysins in Listeria monocytogenes bacteriophages: a  
 RT new class of enzymes and evidence for conserved holin genes within the  
 RT siphoviral lysis cassettes.";  
 RL Mol. Microbiol. 16:1231-1241 (1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.



RC B97, cv. C1187-2, cv. CML254, cv. CML258, cv. D940Y, cv. I205, cv.  
RC IDS28, cv. IL101, cv. K121, cv. K13, cv. M162W, cv. M017, cv. N28HT,  
RC CV. NC260, cv. NC348, cv. OH43, cv. PA91, cv. T232, cv. TX601, and cv.  
RC W153R;  
RC MEDLINE=22247734; PubMed=12244216;  
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;  
RT "Genetic diversity and selection in the maize starch pathway";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. A272, cv. A6, cv. B103, cv. B14A, cv. B37, cv. B73, cv.  
RC B97, cv. C1187-2, cv. CML254, cv. CML258, cv. D940Y, cv. I205, cv.  
RC IDS28, cv. IL101, cv. K121, cv. K13, cv. M162W, cv. M017, cv. N28HT,  
RC CV. NC260, cv. NC348, cv. OH43, cv. PA91, cv. T232, cv. TX601, and cv.  
RC W153R;  
RC Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV;  
RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AY235251; AAP03728.1; -  
DR EMBL; AY235252; AAP03729.1; -  
DR EMBL; AY235253; AAP03730.1; -  
DR EMBL; AY235254; AAP03731.1; -  
DR EMBL; AY235255; AAP03732.1; -  
DR EMBL; AY235256; AAP03733.1; -  
DR EMBL; AY235257; AAP03734.1; -  
DR EMBL; AY235258; AAP03735.1; -  
DR EMBL; AY235259; AAP03736.1; -  
DR EMBL; AY235260; AAP03737.1; -  
DR EMBL; AY235262; AAP03739.1; -  
DR EMBL; AY235265; AAP03742.1; -  
DR EMBL; AY235266; AAP03743.1; -  
DR EMBL; AY235267; AAP03744.1; -  
DR EMBL; AY235268; AAP03745.1; -  
DR EMBL; AY235269; AAP03746.1; -  
DR EMBL; AY235270; AAP03747.1; -  
DR EMBL; AY235271; AAP03748.1; -  
DR EMBL; AY235272; AAP03749.1; -  
DR EMBL; AY235273; AAP03750.1; -  
DR EMBL; AY235274; AAP03751.1; -  
DR EMBL; AY235275; AAP03752.1; -  
DR EMBL; AY235276; AAP03753.1; -  
DR EMBL; AY235277; AAP03754.1; -  
DR EMBL; AY235278; AAP03755.1; -  
DR EMBL; AY235279; AAP03756.1; -  
DR GO; GO:0005507; P:copper ion binding; IEA.  
DR GO; GO:0005489; P:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000923; BlueCu 1.  
DR PROSITE; PS00196; COPPER\_BLUE; 1.  
FT NON TER 139 139  
SQ SEQUENCE 139 AA; 14579 MW; 7BC21F5F025F418D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8  
|:|:|:|:|  
Db 17 PRLTGGG 23

## RESULT 28

Q89GB3 PRELIMINARY; PRT; 181 AA.  
ID Q89GB3  
AC Q89GB3;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE BL6432 protein.  
GN BL6432.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OC NCBI\_TaxID=375;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AF005958; BAC51697.1; -  
KW Complete proteome.  
SQ SEQUENCE 181 AA; 20587 MW; 73C8E7C8781DE6EF CRC64;

Query Match 67.9%; Score 36; DB 16; Length 181;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRVTGGGA 9  
|:|:|:|:|  
Db 70 PRITGGGA 77

## RESULT 29

Q8KVT4 PRELIMINARY; PRT; 222 AA.  
ID Q8KVT4  
AC Q8KVT4;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE EcrF (Fragment).  
GN ECRF.  
OS Streptomyces coelicolor A3(2).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomyces; Streptomycetaceae; Streptomyces.  
ON NCBI\_TaxID=100226;

[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21588047; PubMed=11731481;  
RA Huang J., Lin C.J., Pan K.H., Cohen S.N.;  
RT "Global analysis of growth phase responsive gene expression and  
RT regulation of antibiotic biosynthetic pathways in Streptomyces  
RT coelicolor using DNA microarrays";  
RL Genes Dev. 15:3183-3192(2001).  
DR EMBL; AF425394; AM78437.1; -  
DR GO; GO:0003824; P:catalytic activity; IEA.  
DR GO; GO:0006464; P:protein modification; IEA.  
DR InterPro; IPR004143; BPL\_LipA\_LipB.  
DR Pfam; PF03099; BPL\_LipA\_LipB; 1.  
FT NON TER 1 1

SQ SEQUENCE 222 AA; 24231 MW; A4CID616984440FF CRC64;

Query Match 67.9%; Score 36; DB 2; Length 222;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RVTGGGAM 10  
|:|:|:|:|  
Db 39 RISGGGAM 46

## RESULT 30

Q825D0 PRELIMINARY; PRT; 244 AA.  
ID Q825D0  
AC Q825D0;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN SAV7528.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

```

FT NON_TER 139 139
SQ SEQUENCE 139 AA; 14479 MW; 0D196F5F025F493B CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
Db 17 PRLTGGG 23
||:||||
||:||||

RESULT 26
Q84P36 PRELIMINARY; PRT; 139 AA.
ID Q84P36
AC Q84P36;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Amylose extender 1 (Fragment).
GN AE1.
OS Zea mays (subsp. parviglumis) (Balsas teosinte).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=76912;
ON [1]
RN RN SEQUENCE FROM N.A.
RP RP STRAIN=cv. PI566691;
RX MEDLINE=22247734; PubMed=12244216;
RA Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
[2]
RN RN SEQUENCE FROM N.A.
RP RP STRAIN=cv. PI566691;
RA Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S. IV.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY232420; AA092760.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR NON_TER 139 139
SQ SEQUENCE 139 AA; 14638 MW; 760030BF025D935D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
Db 17 PRLTGGG 23
||:||||
||:||||

RESULT 27
Q84JQ6 PRELIMINARY; PRT; 139 AA.
ID Q84JQ6
AC Q84JQ6;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Amylose extender starch-branching enzyme (Fragment).
GN AE1.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4578;
ON [1]
RN RN SEQUENCE FROM N.A.
RP RP STRAIN=cv. A272, cv. A6, cv. B103, cv. B14A, cv. B37, cv. B73, cv.

```

Q84P38;  
01-JUN-2003 (TReMBLrel. 24, Created)  
01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
Amylose extender 1 (Fragment).  
AE1.  
Zea mays (subsp. parviglumis) (Baleas teosinte).  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
PACCAD clade; Panicoideae; Andropogoneae; Zea.  
NCBI\_TaxID=76912;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. PI566686;  
RX MEDLINE=22247734; PubMed=12244216;  
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;  
RL "Genetic diversity and selection in the maize starch pathway.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. PI566686;  
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY232418; AAC92758.1; -  
GO; GO:0005507; F:copper ion binding; IEA.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000923; BlueCu\_1.  
DR PROSITE; PS00196; COPPER\_BLUE; 1.  
DR NON TPR 139 139  
FT SEQUENCE 139 AA; 14525 MW; 4D00981AD86EC007 CRC64;  
SQ

Query Match 67.9%; Score 36; DB 10; Length 139;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8  
||:||||  
DB 17 PRLTGGG 23

RESULT 25  
Q84P37 PRELIMINARY; PRT; 139 AA.  
AC Q84P37;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Amylose extender 1 (Fragment).  
GN AE1.  
OS Zea mays (subsp. parviglumis) (Baleas teosinte).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=76912;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. PI566688;  
RX MEDLINE=22247734; PubMed=12244216;  
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;  
RL "Genetic diversity and selection in the maize starch pathway.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. PI566688;  
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY232419; AAC92759.1; -  
GO; GO:0005507; F:copper ion binding; IEA.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000923; BlueCu\_1.  
DR PROSITE; PS00196; COPPER\_BLUE; 1.  
DR

```

SQ SEQUENCE 139 AA; 14664 MW; 1E45805F0248DE93 CRC64;
Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 17 PRLTGGG 23

RESULT 21
Q84P41 PRELIMINARY; PRT; 139 AA.
AC Q84P41;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amylose extender 1 (Fragment).
GN AE1.
OS Zea mays (subsp. parviglumis) (Balsas teosinte).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=76912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. BK-5;
RX MEDLINE=22247734; PubMed=12244216;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. BK-5;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. BK-5;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY232415; AAC92755.1; -
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu_1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
FT NON TER 139
SQ SEQUENCE 139 AA; 14736 MW; 780EC15F024C935D CRC64;
Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 17 PRLTGGG 23

RESULT 22
Q84P40 PRELIMINARY; PRT; 139 AA.
AC Q84P40;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amylose extender 1 (Fragment).
GN AE1.
OS Zea mays (subsp. parviglumis) (Balsas teosinte).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=76912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. IC-3;
RX MEDLINE=22247734; PubMed=12244216;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;
RT "Genetic diversity and selection in the maize starch pathway.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. IC-3;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY232416; AAC92756.1; -
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu_1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
FT NON TER 139
SQ SEQUENCE 139 AA; 14494 MW; 7BD905F058AAF03D CRC64;
Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 17 PRLTGGG 23

RESULT 23
Q84P39 PRELIMINARY; PRT; 139 AA.
AC Q84P39;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amylose extender 1 (Fragment).
GN AE1.
OS Zea mays (subsp. parviglumis) (Balsas teosinte).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=76912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. PI331785;
RX MEDLINE=22247734; PubMed=12244216;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. PI331785;
RA Whitt S.R., Wilson L.M., Tenaillon M.I., Gaut B.S., Buckler E.S. IV.;
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY232417; AAC92757.1; -
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu_1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
FT NON TER 139
SQ SEQUENCE 139 AA; 14579 MW; 7BC21F5F025F418D CRC64;
Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRVTGGG 8
Db 17 PRLTGGG 23

RESULT 24
Q84P38 PRELIMINARY; PRT; 139 AA.
ID Q84P38

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=95207408; PubMed=7534933;
RA Laine M.J., Metzler M.C., Zhong Y.P.;
RT "IS1237 a repetitive chromosomal element from Clavibacter xyli subsp.
RT corydonitis is related to insertion sequences from gram-negative and
RT gram-positive bacteria.";
RL Plasmid 32:270-279 (1994).
DR EMBL; X75973; CAA53586.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000531; TonB boxC.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
SQ SEQUENCE 127 AA; 13977 MW; 7A9DD33E922F3DF8 CRC64;

Query Match 67.9%; Score 36; DB 2; Length 127;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
DB 114 TPRIVNGGA 122

RESULT 18
Q84QG3 PRELIMINARY; PRT; 139 AA.
AC Q84QG3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Amylose extender starch-branching enzyme (Fragment).
GN AEI.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=cv. EPI;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RL [2]
RN 2
RP SEQUENCE FROM N.A.
RX STRAIN=cv. EPI;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S. IV.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY235263; AAP03740.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR NON_TER 139 139
SQ SEQUENCE 139 AA; 14494 MW; 7BD905F058AAF03D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 17 PRLTGGG 23

RESULT 20
Q84QG1 PRELIMINARY; PRT; 139 AA.
AC Q84QG1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Amylose extender starch-branching enzyme (Fragment).
GN AEI.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=cv. F2;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RL [2]
RN 2
RP SEQUENCE FROM N.A.
RX STRAIN=cv. F2;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S. IV.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY235264; AAP03741.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR NON_TER 139 139
SQ SEQUENCE 139 AA; 14450 MW; 7BD905EEB678F03D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 17 PRLTGGG 23

RESULT 19
Q84QG2 PRELIMINARY; PRT; 139 AA.
AC Q84QG2;

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Amylose extender starch-branching enzyme (Fragment).
GN AEI.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=cv. EPI;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RL [2]
RN 2
RP SEQUENCE FROM N.A.
RX STRAIN=cv. EPI;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S. IV.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY235263; AAP03740.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR NON_TER 139 139
SQ SEQUENCE 139 AA; 14494 MW; 7BD905F058AAF03D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 17 PRLTGGG 23

RESULT 20
Q84QG1 PRELIMINARY; PRT; 139 AA.
AC Q84QG1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Amylose extender starch-branching enzyme (Fragment).
GN AEI.
OS Zea mays subsp. mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4578;
RN 1
RP SEQUENCE FROM N.A.
RX STRAIN=cv. F2;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S.;
RT "Genetic diversity and selection in the maize starch pathway.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:12959-12962(2002).
RL [2]
RN 2
RP SEQUENCE FROM N.A.
RX STRAIN=cv. F2;
RX MEDLINE=22247734; PubMed=12244216;
RX Whitt S.R., Wilson L.M., Tenailon M.I., Gaut B.S., Buckler E.S. IV.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY235264; AAP03741.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR NON_TER 139 139
SQ SEQUENCE 139 AA; 14450 MW; 7BD905EEB678F03D CRC64;

Query Match 67.9%; Score 36; DB 10; Length 139;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 17 PRLTGGG 23

RESULT 19
Q84QG2 PRELIMINARY; PRT; 139 AA.
AC Q84QG2;

```

OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005041; BAC72891.1;  
 DR GO; GO:0009274; C:cell wall (sensu Bacteria); IEA.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0008658; F:penicillin binding; IEA.  
 DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.  
 DR InterPro; IPR001264; Glyco trans 51.  
 DR InterPro; IPR001460; Transpeptidase.  
 DR Pfam; PF00912; Transglycosyl; 1.  
 DR Pfam; PF00905; Transpeptidase; 1.  
 DR ProDom; PD001895; Glyco\_trans\_51; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 748 AA; 78005 MW; 9073C631F670EEA1 CRC64;  
 Query Match 69.8%; Score 37; DB 16; Length 748;  
 Best Local Similarity 75.0%; Pred. No. 4.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PRVTGGGA 9  
 Db 593 PRINGGGA 600  
 RESULT 15  
 Q82C77 ID Q82C77 PRELIMINARY; PRT; 54 AA.  
 AC Q82C77;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN SAV5477.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005043; BAC73189.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 54 AA; 5884 MW; 7FE1DFB025140016 CRC64;  
 Query Match 67.9%; Score 36; DB 16; Length 54;  
 Best Local Similarity 60.0%; Pred. No. 50;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGAM 10  
 Db 28 SPYMTGGGAL 37  
 RESULT 16  
 Q92ZW2 ID Q92ZW2 PRELIMINARY; PRT; 85 AA.  
 AC Q92ZW2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein RA0336.  
 GN RA0336 OR SMA0638.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OG Plasmid pSymA (megaplasmid 1).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396509; PubMed=11481432;  
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
 RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
 RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
 RT "Nucleotide sequence and predicted functions of the entire  
 RT Sinorhizobium meliloti pSymA megaplasmid.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
 DR EMBL; AE007225; AAK64994.1;  
 DR PIR; H95303; H95303.  
 DR GO; GO:004821; C:extrachromosomal DNA; IEA.  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SQ SEQUENCE 85 AA; 9300 MW; 8C966E072243521C CRC64;  
 Query Match 67.9%; Score 36; DB 16; Length 85;  
 Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPRVTGGGA 9  
 Db 31 TFSVIGGGA 39  
 RESULT 17  
 Q46488 ID Q46488 PRELIMINARY; PRT; 127 AA.  
 AC Q46488;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE ORF2 frame shift at position 228.  
 OS Leifsonia xyl.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Micrococccineae; Microbacteriaceae; Leifsonia.  
 OX NCBI\_TaxID=1575;  
 RN [1]

Query Match 69.8%; Score 37; DB 16; Length 345;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
 Db 192 PKLAGGGAM 200

## RESULT 11

Q9FE35 PRELIMINARY; PRT; 379 AA.  
 AC Q9FE35;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Putative DNA-binding protein homolog.  
 GN P0688A04.2 OR P0006C01.17.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone:P0688A04.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone:P0006C01.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP002839; BAB19096.1; -  
 DR EMBL; AP002744; BAB19075.1; -  
 DR Gramene; Q9FE35; -  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR003657; WRKY.  
 DR Pfam; PF03106; WRKY; 1.  
 DR PROSITE; PSS0811; WRKY; 1.  
 DR DNA-binding.  
 KW KW  
 SQ SEQUENCE 379 AA; 39856 MW; B233AD060113C84D CRC64;

Query Match 69.8%; Score 37; DB 10; Length 379;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
 Db 339 PAVAGGGAM 347

## RESULT 12

Q7XAA2 PRELIMINARY; PRT; 379 AA.  
 AC Q7XAA2;  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE WRKY16.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yao Q., Peng R., Xiong A.;  
 RT "Isolation of rice WRKY protein through W-box bait vector by modified

RT Yeast one-hybrid system method.";  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY341856; AAQ20915.1; -  
 SQ SEQUENCE 379 AA; 39856 MW; B233AD060113C84D CRC64;

Query Match 69.8%; Score 37; DB 10; Length 379;  
 Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
 Db 339 PAVAGGGAM 347

## RESULT 13

Q9ABH5 PRELIMINARY; PRT; 653 AA.  
 AC Q9ABH5;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Penicillin-binding protein, 1A family.  
 GN CC0252  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uitterback T., Iran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).  
 DR EMBL; AS005698; AAK22239.1; -  
 DR PIR; C87280; C87280.  
 DR TIGR; CC0252; -  
 DR GO; GO:0009274; C:cell wall (sensu Bacteria); IEA.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0008658; F:penicillin binding; IEA.  
 DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.  
 DR InterPro; IPR001264; Glyco\_trans\_51.  
 DR InterPro; IPR001460; Transpeptidase.  
 DR Pfam; PF00912; Transglycosyl; 1.  
 DR Pfam; PF00905; Transpeptidase; 1.  
 DR ProDom; PD001895; Glyco\_trans\_51; 1.  
 DR Complete proteome.  
 KW KW  
 SQ SEQUENCE 653 AA; 70503 MW; 395D98388A3772DF CRC64;

Query Match 69.8%; Score 37; DB 16; Length 653;  
 Best Local Similarity 81.8%; Pred. No. 3.9e+02;  
 Matches 9; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 TP--RVTGGGA 9  
 Db 576 TPMKRVGGGA 586

## RESULT 14

Q82D09 PRELIMINARY; PRT; 748 AA.  
 AC Q82D09;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Putative penicillin-binding protein.  
 GN PBP9 OR SAV179.

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative lipopate-protein ligase A.  
 GN BPP2921.  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640432; CAE38214.1; -.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 258 AA; 28296 MW; 2246BDF651E190AF CRC64;

Query Match 69.8%; Score 37; DB 16; Length 258;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 RVTGGGAM 10  
 ||:|||||  
 Db 81 RVSGGGAM 88

RESULT 8  
 Q7VYMI PRELIMINARY; PRT; 258 AA.  
 AC Q7VYMI;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative lipopate-protein ligase A.  
 GN BPI297.  
 OS Bordetella pertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=520;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebathia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640414; CAE41593.1; -.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 258 AA; 28296 MW; 2246BDF651E190AF CRC64;

Query Match 69.8%; Score 37; DB 16; Length 258;  
 Best Local Similarity 87.5%; Pred. No. 1.6e+02;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RVTGGGAM 10  
 ||:|||||  
 Db 81 RVSGGGAM 88

RESULT 9  
 Q7V426 PRELIMINARY; PRT; 264 AA.  
 AC Q7V426;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein precursor.  
 GN PWT2147.  
 OS Prochlorococcus marinus (strain MIT 9313).  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=74547;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22825698; PubMed=12917642;  
 RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,  
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
 RA Webb E.A., Zinser E.R., Chisholm S.W.;  
 RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
 RT niche differentiation.";  
 RL Nature 424:1042-1047(2003).  
 DR EMBL; BX572101; CAE22321.1; -.  
 KW Hypothetical protein; Signal; Complete proteome.  
 FT SIGNAL 1 32  
 SQ SEQUENCE 264 AA; 28534 MW; 4E3D89A3750C51D0 CRC64;

Query Match 69.8%; Score 37; DB 16; Length 264;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTCGG 8  
 ||:|||||  
 Db 253 TPRVTCGG 260

RESULT 10  
 Q7VG46 PRELIMINARY; PRT; 345 AA.  
 AC Q7VG46;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN HH1478.  
 OS Helicobacter hepaticus.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=32025;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 51449 / 3B1;  
 RX MEDLINE=22709201; PubMed=12810954;  
 RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,  
 RA Bell M., Droege K., Koenig J., Fischer H.P., Ge Z., Hoerster A.,  
 RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,  
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;  
 RT "The complete genome sequence of the carcinogenic bacterium  
 RT Helicobacter hepaticus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).  
 DR EMBL; AB017148; AAP78075.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 345 AA; 38138 MW; 05F9948CDE1880A5 CRC64;



```

DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 340 AA; 36503 MW; 3C2540E9B3A81C8A CRC64;

Query Match 71.7%; Score 38; DB 11; Length 340;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 43 PRITGGG 49

RESULT 5
ID Q9SYG2 PRELIMINARY; PRT; 383 AA.
AC Q9SYG2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE F1511.14 protein (Putative DNA-binding protein) (Hypothetical protein)
GN F1511.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ccv. Columbia;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S., Li J., Kremenetskaia I., Luros J., Ngan I., Gonzalez A., Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RA "Arabidopsis thaliana chromosome 1 BAC F1511 sequence.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K., Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RA "Full length cDNA of gene F1511.14 (GI:4587547).";
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RA "Full length cDNA of gene F1511.14 (GI:4587547).";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N., Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RA "Full-length messenger RNA sequences greatly improve genome annotation.";
RT Genome Biol. 0:0-0(2002).
RL

[5]
RN SEQUENCE FROM N.A.
RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R., RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006577; AAD25778.1; -
DR EMBL; AF360325; AAK26035.1; -
DR EMBL; AY056333; AAL07182.1; -
DR EMBL; AY084772; AAM61340.1; -
DR PIR; C96581; C96581.
DR GO; GO:0003677; F-DNA binding; IEA.
DR Hypothetical protein; DNA-binding.
KW SEQUENCE 383 AA; 41731 MW; D7AE4D93DAF2B1EC CRC64;

Query Match 71.7%; Score 38; DB 10; Length 383;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
DB 62 TPSVTGGG 69

RESULT 6
ID Q7WIG0 PRELIMINARY; PRT; 258 AA.
AC Q7WIG0;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative lipote-protein ligase A.
GN BB2891.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N., Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A., Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I., Chillingworth T., Collins M., Cronin A., Davis P., Doggett J., Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagers K., Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C., Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640445; CAE33383.1; -
KW Ligase; Complete proteome.
SQ SEQUENCE 258 AA; 28296 MW; 2246BDF651E190AF CRC64;

Query Match 69.8%; Score 37; DB 16; Length 258;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 RVTGGGAM 10
DB 81 RVSGGGAM 88

RESULT 7
ID Q7W618 PRELIMINARY; PRT; 258 AA.
AC Q7W618;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

```



```

RP SEQUENCE FROM N.A.
RC STRAIN=129/SvU, and Swiss; TISSUE=Lung;
RA Verghese G.M., Caughey G.H.;
RT "Molecular cloning and characterization of mouse prostatic, a type I
RT membrane-associated serine protease of the gamma-tryptase/prostasin
RT gene family.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

SEQUENCE FROM N.A.
RP Kitamura K., Takefumi N., Kimio T.;
RA "mouse serine protease.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC EMBL; BC003851; AA030851.1; -
DR EMBL; AF378086; AAL063320.1; -
DR EMBL; AF378085; AAL063319.1; -
DR EMBL; AF038244; BAB82496.1; -
DR HSSP; P00734; IUVS.
DR MGD; MGI:1923810; Prss8.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Protease; Serine protease.
KW HYDROLASE; PROTEASE; SERINE PROTEASE.
SQ SEQUENCE 339 AA; 36216 MW; BC2D888BC057AF10 CRC64;

Query Match 73.6%; Score 39; DB 11; Length 339;
Best Local Similarity 75.0%; Pred. No. 92;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps

QY 2 PRVTGGGA 9
DB 43 PRITGGGS 50
:::|:::

RESULT 4
ID Q8BJV6 PRELIMINARY; PRT; 340 AA.
OC AC Q8BJV6;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Prostasin.
GN PRSS8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Eye;
RC MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK078696; BAC37362.1; -
DR MGD; MGI:1923810; Prss8.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.

```

382	32	60.4	131	2	Q9F1V4	Q9F1V4 streptomyc	455	32	60.4	269	2	Q9AXK4	Q9AXK4 legionella
383	32	60.4	134	11	Q8C504	Q8C504 mus musculus	456	32	60.4	271	4	Q8ND25	Q8ND25 homo sapien
384	32	60.4	135	3	Q876F7	Q876F7 saccharomyc	457	32	60.4	271	5	Q1B487	Q1B487 penaeus van
385	32	60.4	135	3	Q876P6	Q876P6 saccharomyc	458	32	60.4	271	16	Q7U379	Q7U379 bordetella
386	32	60.4	135	3	Q875W0	Q875W0 saccharomyc	459	32	60.4	271	16	Q7U368	Q7U368 bordetella
387	32	60.4	135	3	Q875V9	Q875V9 saccharomyc	460	32	60.4	272	2	Q68500	Q68500 streptomyc
388	32	60.4	141	16	Q8CX50	Q8CX50 leptospira	461	32	60.4	279	12	Q68458	Q68458 hepatocis c
389	32	60.4	143	3	Q87503	Q87503 saccharomyc	462	32	60.4	281	2	Q9EY14	Q9EY14 streptomyc
390	32	60.4	154	3	Q87503	Q87503 saccharomyc	463	32	60.4	287	16	Q87P21	Q87P21 vibrio para
391	32	60.4	155	5	Q9VZK6	Q9VZK6 caulobacter	464	32	60.4	294	2	Q44382	Q44382 agrobacteri
392	32	60.4	164	10	Q7XL96	Q7XL96 oryza sativ	465	32	60.4	294	9	Q9AYZ8	Q9AYZ8 bacterioph
393	32	60.4	174	10	Q8H586	Q8H586 oryza sativ	466	32	60.4	295	16	Q985T2	Q985T2 rhizobium l
394	32	60.4	175	16	Q50451	Q50451 mycobacteri	467	32	60.4	309	16	Q8YC55	Q8YC55 brucella me
395	32	60.4	175	16	Q7U088	Q7U088 mycobacteri	468	32	60.4	309	16	Q8FW71	Q8FW71 brucella su
396	32	60.4	176	11	Q9D367	Q9D367 mus musculu	469	32	60.4	316	16	Q8XXK4	Q8XXK4 ralstonia s
397	32	60.4	177	5	Q9GV64	Q9GV64 drosophila	470	32	60.4	328	16	Q99V65	Q99V65 staphylococ
398	32	60.4	177	5	Q9GV56	Q9GV56 drosophila	471	32	60.4	328	16	Q8NKB5	Q8NKB5 staphylococ
399	32	60.4	177	5	Q9GV66	Q9GV66 drosophila	472	32	60.4	328	16	Q8CT48	Q8CT48 staphylococ
400	32	60.4	177	5	Q9GV63	Q9GV63 drosophila	473	32	60.4	329	16	Q97QP1	Q97QP1 streptococ
401	32	60.4	177	5	Q9GV60	Q9GV60 drosophila	474	32	60.4	329	16	Q8DPR1	Q8DPR1 streptococ
402	32	60.4	177	5	Q9GV57	Q9GV57 drosophila	475	32	60.4	331	2	Q9ZJ09	Q9ZJ09 rhodococ
403	32	60.4	177	5	Q9GNAS	Q9GNAS drosophila	476	32	60.4	333	2	Q32631	Q32631 helicobacte
404	32	60.4	177	5	Q9GV68	Q9GV68 drosophila	477	32	60.4	333	4	Q9H506	Q9H506 homo sapien
405	32	60.4	177	5	Q9GN47	Q9GN47 drosophila	478	32	60.4	336	16	Q88U17	Q88U17 lactobacill
406	32	60.4	177	5	Q9GV61	Q9GV61 drosophila	479	32	60.4	337	10	Q9F1W5	Q9F1W5 arabidopsis
407	32	60.4	177	5	Q9GV67	Q9GV67 drosophila	480	32	60.4	345	16	Q98RH7	Q98RH7 mycoplasma
408	32	60.4	177	5	Q7YMU7	Q7YMU7 caenorhabdi	481	32	60.4	354	16	Q7UWW8	Q7UWW8 rhodopirell
409	32	60.4	181	5	Q9GV59	Q9GV59 drosophila	482	32	60.4	362	5	Q9U181	Q9U181 leishmania
410	32	60.4	181	5	Q9GV62	Q9GV62 drosophila	483	32	60.4	363	17	Q28258	Q28258 archaeoglob
411	32	60.4	181	5	Q9GND5	Q9GND5 drosophila	484	32	60.4	364	12	Q91B45	Q91B45 pan rhadino
412	32	60.4	181	5	Q9GV65	Q9GV65 drosophila	485	32	60.4	378	16	F74348	F74348 synechocyst
413	32	60.4	181	5	Q9GV58	Q9GV58 drosophila	486	32	60.4	379	16	Q8YMU7	Q8YMU7 anabaena sp
414	32	60.4	184	16	Q88JY6	Q88JY6 pseudomonas	487	32	60.4	384	10	Q8LQ65	Q8LQ65 oryza sativ
415	32	60.4	191	12	Q8JPC7	Q8JPC7 hepatitis c	488	32	60.4	385	2	Q3EX94	Q3EX94 pseudomonas
416	32	60.4	192	16	Q92MB6	Q92MB6 rhizobium s	489	32	60.4	389	12	Q91B44	Q91B44 pan rhadino
417	32	60.4	193	16	Q8XTT3	Q8XTT3 ralstonia s	490	32	60.4	389	12	Q91B43	Q91B43 pan rhadino
418	32	60.4	199	5	Q9N3U3	Q9N3U3 caenorhabdi	491	32	60.4	390	17	Q96Y06	Q96Y06 sulfolobus
419	32	60.4	200	2	Q9K4Q7	Q9K4Q7 salmonella	492	32	60.4	396	3	Q870Y6	Q870Y6 neurospora
420	32	60.4	203	16	Q89JG1	Q89JG1 bradyrhizob	493	32	60.4	396	16	Q7VQG3	Q7VQG3 helicobacte
421	32	60.4	217	16	Q86622	Q86622 aquifex aeo	494	32	60.4	398	16	Q9Z678	Q9Z678 streptomyc
422	32	60.4	218	10	Q8LJW4	Q8LJW4 sorghum bic	495	32	60.4	403	16	Q8NTE5	Q8NTE5 corynebacte
423	32	60.4	220	16	Q82CY7	Q82CY7 streptomyc	496	32	60.4	408	5	Q8IEG9	Q8IEG9 plasmodium
424	32	60.4	222	16	Q9AC9C	Q9AC9C caulobacter	497	32	60.4	413	16	Q82UG8	Q82UG8 streptomyc
425	32	60.4	227	4	Q9H083	Q9H083 homo sapien	498	32	60.4	417	16	Q9KSF5	Q9KSF5 vibrio chol
426	32	60.4	227	11	Q91V17	Q91V17 mus musculu	499	32	60.4	417	16	Q8DAE4	Q8DAE4 vibrio vuln
427	32	60.4	228	16	Q89MB3	Q89MB3 bradyrhizob	500	32	60.4	417	16	Q87N19	Q87N19 vibrio para
428	32	60.4	230	16	Q8FDM5	Q8FDM5 escherichia							
429	32	60.4	230	16	Q83JMO	Q83JMO shigella fl							
430	32	60.4	234	10	Q94H31	Q94H31 oryza sativ							
431	32	60.4	235	10	Q9LQ73	Q9LQ73 arabidopsis							
432	32	60.4	235	10	Q9LQ73	Q9LQ73 lycopersico							
433	32	60.4	236	10	Q80845	Q80845 arabidopsis							
434	32	60.4	236	16	Q8PIE3	Q8PIE3 xanthomonas							
435	32	60.4	236	16	Q8F729	Q8F729 xanthomonas							
436	32	60.4	238	5	Q9GQH5	Q9GQH5 drosophila							
437	32	60.4	239	16	Q9RV15	Q9RV15 deinococcus							
438	32	60.4	239	16	Q8NU17	Q8NU17 corynebacte							
439	32	60.4	240	10	Q9FW16	Q9FW16 oryza sativ							
440	32	60.4	240	10	Q7KX9	Q7KX9 oryza sativ							
441	32	60.4	242	17	Q9HPH6	Q9HPH6 halobacteri							
442	32	60.4	243	5	Q9GQH4	Q9GQH4 drosophila							
443	32	60.4	247	16	Q9RR93	Q9RR93 deinococcus							
444	32	60.4	249	15	Q86342	Q86342 bovine foam							
445	32	60.4	249	15	Q8ALT9	Q8ALT9 bovine foam							
446	32	60.4	252	5	Q9GQI0	Q9GQI0 drosophila							
447	32	60.4	252	10	Q9LWV4	Q9LWV4 oryza sativ							
448	32	60.4	257	16	Q9ZMA1	Q9ZMA1 rhizobium m							
449	32	60.4	258	16	Q8Y4V6	Q8Y4V6 listeria mo							
450	32	60.4	260	2	Q9F5G8	Q9F5G8 agrobacteri							
451	32	60.4	262	17	Q979Y9	Q979Y9 thermoplaem							
452	32	60.4	264	17	Q97U62	Q97U62 sulfolobus							
453	32	60.4	269	2	Q9AQN5	Q9AQN5 pseudomonas							
454	32	60.4	269	2	Q8GI30	Q8GI30 pseudomonas							

## ALIGNMENTS

## RESULT 1

ID	Q7TER5	PRELIMINARY;	PRT;	561 AA.
AC	Q7TER5			
DT	01-OCT-2003 (TReMBLrel. 25, Created)			
DT	01-OCT-2003 (TReMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)			
DE	Mutant UL83.			
OS	Human cytomegalovirus.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
OX	NCBI_TaxID=10359;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Han J., Li Y., Wang M.;			
RT	"Construction of human cytomegalovirus cDNA library and screening of			
RT	pp65 clone."			
RT	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY301013; AAP59842.1; "			
SQ	SEQUENCE 561 AA; 62872 MW; 2B322EB449FD9A30 CRC64;			

Query Match 100.0%; Score 53; DB 12; Length 561;

236	33	62.3	219	4	Q9NPE2	Q9ape2 homo sapien	309	33	62.3	409	16	Q7U5P8	Q7u5p8 synechococc
237	33	62.3	223	16	Q7U471	Q7u471 synechococc	310	33	62.3	411	5	Q8I431	Q8i431 plasmodium
238	33	62.3	225	16	Q92S74	Q92st4 rhizobium m	311	33	62.3	434	10	Q8LSI4	Q8lsi4 chlamydomon
239	33	62.3	228	16	Q82AB8	Q82ab8 streptomyce	312	33	62.3	445	16	Q9PKY9	Q9pky9 chlamydia m
240	33	62.3	229	16	Q8ZJV3	Q8zjv3 salmonella	313	33	62.3	448	10	Q94IU2	Q94iu2 oryza sativ
241	33	62.3	230	11	Q9EQY9	Q9eqy9 mus musculu	314	33	62.3	464	16	Q7WGS4	Q7wgs4 bordetella
242	33	62.3	231	5	Q15714	Q15714 dictyosteli	315	33	62.3	464	16	Q7W4M8	Q7w4m8 bordetella
243	33	62.3	237	10	Q8S7Q1	Q8s7q1 oryza sativ	316	33	62.3	471	2	Q9S946	Q9s946 morganella
244	33	62.3	242	16	Q938N5	Q938n5 streptococc	317	33	62.3	474	17	Q8PTL1	Q8ptl1 methanosarc
245	33	62.3	247	16	Q7WH87	Q7wh87 bordetella	318	33	62.3	475	16	Q9A1N9	Q9aln9 streptococc
246	33	62.3	247	16	Q7W9H5	Q7w9h5 bordetella	319	33	62.3	475	2	Q8RIQ0	Q8riq0 streptomyce
247	33	62.3	247	16	Q7W9N1	Q7wn1 bordetella	320	33	62.3	476	2	Q8P2T1	Q8p2t1 streptococc
248	33	62.3	248	16	Q89P10	Q89p10 bradyrhizob	321	33	62.3	477	16	Q8K8S0	Q8k8s0 streptococc
249	33	62.3	249	9	Q80074	Q80074 staphylococ	322	33	62.3	479	16	Q8E3A5	Q8e3a5 streptococc
250	33	62.3	249	9	Q8SDM9	Q8sdm9 staphylococ	323	33	62.3	479	16	Q8DXN5	Q8dxn5 streptococc
251	33	62.3	249	16	Q8NMV6	Q8nmv6 staphylococ	324	33	62.3	485	16	Q97N19	Q97ni9 streptococc
252	33	62.3	249	17	Q59156	Q59156 pyrococcus	325	33	62.3	485	16	Q93DB1	Q93db1 streptococc
253	33	62.3	249	17	Q9V0V8	Q9v0v8 pyrococcus	326	33	62.3	485	16	Q82AZ3	Q82az3 streptomyce
254	33	62.3	250	9	Q9G034	Q9g034 bacterioph	327	33	62.3	487	5	Q86CS3	Q86cs3 drosophila
255	33	62.3	250	16	Q99SP8	Q99sp8 staphylococ	328	33	62.3	488	5	Q18545	Q18545 caenorhabdi
256	33	62.3	251	16	Q931J4	Q931j4 staphylococ	329	33	62.3	499	10	Q8S7G6	Q8s7g6 oryza sativ
257	33	62.3	254	4	Q96EZ6	Q96ez6 homo sapien	330	33	62.3	508	16	Q8DN94	Q8dn94 streptococc
258	33	62.3	259	17	Q8U153	Q8u153 pyrococcus	331	33	62.3	541	10	Q7XNA0	Q7xna0 oryza sativ
259	33	62.3	260	12	Q7TFN5	Q7tfns rhesus cyto	332	33	62.3	549	10	Q7XPT7	Q7xpt7 oryza sativ
260	33	62.3	265	9	Q9WBTO	Q9wbt0 staphylococ	333	33	62.3	559	11	Q9QW71	Q9qw71 rattus sp.
261	33	62.3	265	16	Q8FNT3	Q8fnt3 corynebacte	334	33	62.3	586	13	Q8JW11	Q8jim1 brachydanio
262	33	62.3	267	10	Q943T4	Q943t4 oryza sativ	335	33	62.3	586	13	Q7ZV12	Q7zvi2 brachydanio
263	33	62.3	288	10	Q7XUS2	Q7xus2 oryza sativ	336	33	62.3	661	16	Q987J6	Q987j6 rhizobium l
264	33	62.3	291	5	Q9VG03	Q9vg03 drosophila	337	33	62.3	665	12	Q91EVA	Q91ev4 cydia pomon
265	33	62.3	292	10	Q7XL79	Q7xl79 oryza sativ	338	33	62.3	691	5	Q8MSV9	Q8msv9 drosophila
266	33	62.3	296	5	Q9V7H9	Q9v7h9 drosophila	339	33	62.3	694	16	Q9L179	Q9ll79 streptomyce
267	33	62.3	305	10	Q8LDL1	Q8ldl1 arabidopsis	340	33	62.3	703	5	Q9VHV7	Q9vhv7 drosophila
268	33	62.3	310	5	Q96649	Q96649 trypanosoma	341	33	62.3	704	10	Q9SB64	Q9sb64 arabidopsis
269	33	62.3	310	5	Q97466	Q97466 trypanosoma	342	33	62.3	706	17	Q9YAG6	Q9yag6 aeropyrum p
270	33	62.3	311	10	Q7X635	Q7x635 oryza sativ	343	33	62.3	733	10	Q7XKB8	Q7xkb8 oryza sativ
271	33	62.3	314	10	Q9M9A3	Q9msa3 arabidopsis	344	33	62.3	818	16	Q8E1H2	Q8eih2 shewanella
272	33	62.3	314	16	Q98PA9	Q98pa9 rhizobium l	345	33	62.3	1003	16	Q8FM34	Q8fm34 corynebacte
273	33	62.3	321	2	P74937	P74937 thermoactin	346	33	62.3	1018	9	Q9ZXE7	Q9zxe7 bacterioph
274	33	62.3	322	16	Q9KAV3	Q9kav3 bacillus ha	347	33	62.3	1023	15	Q9Q0A2	Q9q0a2 chimpanzee
275	33	62.3	329	10	Q9SDG0	Q9sdg0 oryza sativ	348	33	62.3	1023	15	Q9Q094	Q9q094 chimpanzee
276	33	62.3	333	16	Q9CJD4	Q9cjd4 lactococcus	349	33	62.3	1031	15	Q90273	Q90273 chimpanzee
277	33	62.3	335	5	Q8SRC2	Q8src2 encephalico	350	33	62.3	1035	15	Q9Q086	Q9q086 chimpanzee
278	33	62.3	337	16	Q88U78	Q88u78 lactobacill	351	33	62.3	1048	4	Q9R2H5	Q9y2h5 homo sapien
279	33	62.3	337	16	Q938C1	Q938c1 enterococcu	352	33	62.3	1111	16	Q9L0P1	Q9l0p1 streptomyce
280	33	62.3	338	16	Q892M2	Q892m2 clostridium	353	33	62.3	1173	11	Q7TQG1	Q7tqg1 mus musculu
281	33	62.3	339	10	Q949T9	Q949t9 arabidopsis	354	33	62.3	1218	5	Q8IU42	Q8iu42 dictyosteli
282	33	62.3	339	16	Q8E3H3	Q8e3h3 streptococ	355	33	62.3	1264	5	Q9V4U4	Q9v4u4 drosophila
283	33	62.3	339	16	Q8DXV5	Q8dxv5 streptococ	356	33	62.3	1314	5	Q8I172	Q8i172 drosophila
284	33	62.3	346	5	Q9N3W1	Q9n3w1 caenorhabdi	357	33	62.3	1371	16	Q8XQ42	Q8xq42 raistonia s
285	33	62.3	347	10	Q9LY92	Q9ly92 arabidopsis	358	33	62.3	1400	2	Q9LCU0	Q9lcu0 mycobacteri
286	33	62.3	350	16	Q89N33	Q89nx3 bradyrhizob	359	33	62.3	1521	5	Q9VZS2	Q9vzs2 drosophila
287	33	62.3	356	2	Q93TG5	Q93tg5 brucella me	360	33	62.3	1522	5	Q7YSM9	Q7ysm9 trypanosoma
288	33	62.3	358	16	Q8F206	Q8f206 brucella su	361	33	62.3	1531	5	Q9I7T7	Q9i7t7 drosophila
289	33	62.3	358	16	Q7V567	Q7v567 prochloroco	362	33	62.3	1582	2	Q83WE7	Q83we7 micromonosp
290	33	62.3	363	10	Q9XFFV3	Q9xfv3 chlamydomon	363	33	62.3	1683	11	Q7TP67	Q7tp67 rattus norv
291	33	62.3	366	2	Q84HK9	Q84hk9 streptomyce	364	33	62.3	1764	2	Q93T34	Q93t34 haemophilus
292	33	62.3	375	10	Q9FT42	Q9ft42 arabidopsis	365	33	62.3	1843	5	Q9V1S1	Q9v1s1 drosophila
293	33	62.3	375	13	Q7ZYR6	Q7zyr6 xenopus lae	366	33	62.3	2141	5	Q869H2	Q869h2 lymnaea sta
294	33	62.3	375	13	Q7ZW95	Q7zw95 brachydanio	367	33	62.3	2977	5	Q9VAP9	Q9vap9 drosophila
295	33	62.3	376	16	Q829L2	Q829l2 streptomyce	368	33	62.3	3649	2	Q83WE8	Q83we8 micromonosp
296	33	62.3	381	10	Q8LEB2	Q8leb2 arabidopsis	369	33	62.3	4106	16	Q8XQP2	Q8xqp2 raistonia s
297	33	62.3	381	10	Q8RV10	Q8rv10 arabidopsis	370	33	62.3	4290	2	Q9WXC0	Q9wxc0 micromonosp
298	33	62.3	385	13	Q90YW7	Q90yw7 ictalurus p	371	33	62.3	4307	2	Q83WF0	Q83wf0 micromonosp
299	33	62.3	396	13	Q7ZY88	Q7zy88 xenopus lae	372	32.5	61.3	293	10	Q7XTP8	Q7xtp8 oryza sativ
300	33	62.3	401	5	Q917H6	Q9i7h6 drosophila	373	32	60.4	47	15	Q87267	Q87267 chimpanzee
301	33	62.3	401	13	Q7SY96	Q7sy96 xenopus lae	374	32	60.4	58	5	Q84899	Q84899 penaeus van
302	33	62.3	402	16	Q82JF4	Q82jf4 streptomyce	375	32	60.4	82	12	Q9Q3P5	Q9q3ps heparitis c
303	33	62.3	404	10	Q7Y115	Q7y115 oryza sativ	376	32	60.4	89	2	Q9VMA7	Q9vma7 rhizobium e
304	33	62.3	404	16	Q67254	Q67254 aquifex aeo	377	32	60.4	105	10	Q93W83	Q93w83 arabidopsis
305	33	62.3	405	4	Q9GE86	Q96e86 homo sapien	378	32	60.4	115	10	Q8S025	Q8s025 oryza sativ
306	33	62.3	406	10	Q8LA93	Q8la93 arabidopsis	379	32	60.4	126	11	Q8C3N8	Q8c3n8 mus musculu
307	33	62.3	406	17	Q8U209	Q8u209 pyrococcus	380	32	60.4	128	16	Q8PFX7	Q8pfx7 xanthomonas
308	33	62.3	407	10	Q8GYH3	Q8gyh3 arabidopsis	381	32	60.4	129	9	Q9MBI5	Q9mbi5 bacterioph

90	34	64.2	46	12	092572	092572 hepatitis c	163	34	64.2	1063	10	07XS39	07XS39 oryza sativ
91	34	64.2	146	9	07Y5J6	07Y5J6 xanthomonas	164	34	64.2	1076	10	094182	094182 oryza sativ
92	34	64.2	212	10	07XK55	07XK55 oryza sativ	165	34	64.2	1076	10	08LIT2	08LIT2 oryza sativ
93	34	64.2	252	16	09KD35	09KD35 bacillus ha	166	34	64.2	1076	10	07XL89	07XL89 oryza sativ
94	34	64.2	261	16	087DR2	087DR2 xylella fas	167	34	64.2	1081	10	07XL82	07XL82 oryza sativ
95	34	64.2	263	16	08NRH5	08NRH5 corynebacte	168	34	64.2	1088	10	08LMS4	08LMS4 oryza sativ
96	34	64.2	267	17	09HRCS	09HRCS halobacteri	169	34	64.2	1089	10	08SB79	08SB79 oryza sativ
97	34	64.2	273	9	09ZXM4	09ZXM4 bacterioph	170	34	64.2	1090	10	085017	085017 oryza sativ
98	34	64.2	276	2	030879	030879 pseudomonas	171	34	64.2	1109	10	094HM6	094HM6 oryza sativ
99	34	64.2	288	10	08LAI39	08LAI39 arabisdopsis	172	34	64.2	1109	10	07XG30	07XG30 oryza sativ
100	34	64.2	288	10	094960	094960 arabisdopsis	173	34	64.2	1110	5	09BJ49	09BJ49 leishmania
101	34	64.2	298	16	09RV70	09RV70 deinococcus	174	34	64.2	1110	10	08S795	08S795 oryza sativ
102	34	64.2	306	2	085773	085773 rhizobium l	175	34	64.2	1110	10	08RZG2	08RZG2 oryza sativ
103	34	64.2	346	4	095274	095274 homo sapien	176	34	64.2	1110	10	08RZ24	08RZ24 oryza sativ
104	34	64.2	346	4	09UI74	09UI74 homo sapien	177	34	64.2	1110	10	08GSY2	08GSY2 oryza sativ
105	34	64.2	354	16	07ITU7	07ITU7 synecococc	178	34	64.2	1110	10	08LLY6	08LLY6 oryza sativ
106	34	64.2	355	2	09AGV5	09AGV5 corynebacte	179	34	64.2	1110	10	08H719	08H719 oryza sativ
107	34	64.2	356	16	08FQJ6	08FQJ6 corynebacte	180	34	64.2	1110	10	07XX60	07XX60 oryza sativ
108	34	64.2	357	16	088FE9	088FE9 pseudomonas	181	34	64.2	1110	10	07X817	07X817 oryza sativ
109	34	64.2	357	16	087YS4	087YS4 pseudomonas	182	34	64.2	1110	10	07X722	07X722 oryza sativ
110	34	64.2	370	16	08UI65	08UI65 agrobacteri	183	34	64.2	1110	5	Q17346	Q17346 caenorhabdi
111	34	64.2	379	12	09DSP2	09DSP2 saint croix	184	34	64.2	1273	10	084Q03	084Q03 oryza sativ
112	34	64.2	379	16	086306	086306 bifidobacte	185	34	64.2	1352	10	094EB4	094EB4 oryza sativ
113	34	64.2	380	16	097RE2	097RE2 streptococc	186	34	64.2	1359	5	09NKQ3	09NKQ3 leishmania
114	34	64.2	380	16	08DQ93	08DQ93 streptococc	187	34	64.2	1389	13	090269	090269 brachydanio
115	34	64.2	382	2	084HN9	084HN9 streptomyc	188	34	64.2	1419	13	098SW3	098SW3 brachydanio
116	34	64.2	394	16	089N16	089N16 bradyrhizob	189	34	64.2	1548	5	09NE01	09NE01 leishmania
117	34	64.2	403	16	08KFX2	08KFX2 chlorobium	190	34	64.2	1640	10	084R06	084R06 oryza sativ
118	34	64.2	409	17	097UK3	097UK3 sulfolobus	191	34	64.2	1657	10	084MY2	084MY2 oryza sativ
119	34	64.2	440	10	09LGL14	09LGL14 arabisdopsis	192	34	64.2	1758	10	07XU25	07XU25 oryza sativ
120	34	64.2	484	10	08W4R3	08W4R3 arabisdopsis	193	34	64.2	1873	10	07X816	07X816 oryza sativ
121	34	64.2	516	11	08CCC4	08CCC4 mus musculu	194	34	64.2	1882	10	07XNN5	07XNN5 oryza sativ
122	34	64.2	521	16	08RI84	08RI84 fusobacteri	195	34	64.2	1973	10	07XT50	07XT50 oryza sativ
123	34	64.2	523	16	0987K6	0987K6 rhizobium l	196	34	64.2	2525	10	07XP85	07XP85 oryza sativ
124	34	64.2	564	10	09FML1	09FML1 arabisdopsis	197	34	64.2	2535	10	0948C7	0948C7 oryza sativ
125	34	64.2	591	10	07XTZ0	07XTZ0 oryza sativ	198	34	64.2	2535	10	07XG72	07XG72 oryza sativ
126	34	64.2	637	10	08H834	08H834 oryza sativ	199	34	64.2	2877	10	07XEP7	07XEP7 oryza sativ
127	34	64.2	659	10	094L64	094L64 oryza sativ	200	34	64.2	3275	16	08VKM3	08VKM3 mycobacteri
128	34	64.2	661	4	095364	095364 homo sapien	201	34	64.2	3300	16	066304	066304 mycobacteri
129	34	64.2	661	4	08IYI4	08IYI4 homo sapien	202	34	64.2	3507	16	07U270	07U270 mycobacteri
130	34	64.2	666	4	Q13470	Q13470 homo sapien	203	33	62.3	27	12	08QJK6	08QJK6 hepatitis c
131	34	64.2	666	11	099ML2	099ML2 mus musculu	204	33	62.3	27	12	08QJK4	08QJK4 hepatitis c
132	34	64.2	666	11	08K0X9	08K0X9 mus musculu	205	33	62.3	27	12	08QJL4	08QJL4 hepatitis c
133	34	64.2	668	10	09FRF1	09FRF1 oryza sativ	206	33	62.3	27	12	08QJMO	08QJMO hepatitis c
134	34	64.2	689	2	093FZ1	093FZ1 pectobacter	207	33	62.3	27	12	08QJMO	08QJMO hepatitis c
135	34	64.2	690	10	07XRR9	07XRR9 oryza sativ	208	33	62.3	27	12	08QJL9	08QJL9 hepatitis c
136	34	64.2	691	10	07XJX3	07XJX3 oryza sativ	209	33	62.3	27	12	08QJL7	08QJL7 hepatitis c
137	34	64.2	715	10	07XPF6	07XPF6 oryza sativ	210	33	62.3	27	12	08QJMO	08QJMO hepatitis c
138	34	64.2	731	10	07XHR8	07XHR8 oryza sativ	211	33	62.3	47	10	09LE87	09LE87 arabisdopsis
139	34	64.2	740	10	08L8J9	08L8J9 oryza sativ	212	33	62.3	54	12	081859	081859 hepatitis d
140	34	64.2	767	10	07XLM3	07XLM3 oryza sativ	213	33	62.3	57	12	069122	069122 human herpe
141	34	64.2	776	10	08LMK5	08LMK5 oryza sativ	214	33	62.3	57	12	069114	069114 human herpe
142	34	64.2	817	10	07XRF5	07XRF5 oryza sativ	215	33	62.3	87	13	090X98	090X98 measles g
143	34	64.2	860	10	07XK31	07XK31 oryza sativ	216	33	62.3	107	5	081GCG	081GCG xanthophila
144	34	64.2	877	16	08BAW5	08BAW5 pseudomonas	217	33	62.3	111	16	08P6K5	08P6K5 xanthomonas
145	34	64.2	929	10	07XRP3	07XRP3 oryza sativ	218	33	62.3	125	10	09S3J23	09S3J23 arabisdopsis
146	34	64.2	935	10	08W5N8	08W5N8 oryza sativ	219	33	62.3	132	17	08PT50	08PT50 methanosarc
147	34	64.2	935	10	07XDG2	07XDG2 oryza sativ	220	33	62.3	134	13	090W69	090W69 oncorhynch
148	34	64.2	941	3	Q07048	Q07048 baccharomyc	221	33	62.3	141	13	090YH1	090YH1 oncorhynch
149	34	64.2	953	10	07X7D7	07X7D7 oryza sativ	222	33	62.3	147	16	09X8M0	09X8M0 streptomyc
150	34	64.2	959	10	07XW14	07XW14 oryza sativ	223	33	62.3	165	10	08S751	08S751 oryza sativ
151	34	64.2	988	10	084MY8	084MY8 oryza sativ	224	33	62.3	165	10	07XGH6	07XGH6 oryza sativ
152	34	64.2	990	5	09BLV0	09BLV0 leishmania	225	33	62.3	174	16	07VPW5	07VPW5 chlamydia p
153	34	64.2	999	10	09AYJ0	09AYJ0 oryza sativ	226	33	62.3	177	17	08ZTY8	08ZTY8 pyrobaculum
154	34	64.2	999	10	07XHF8	07XHF8 oryza sativ	227	33	62.3	178	2	09KHX7	09KHX7 escherichia
155	34	64.2	1005	10	07X7B1	07X7B1 oryza sativ	228	33	62.3	180	16	08YIV2	08YIV2 brucella me
156	34	64.2	1027	10	084UL6	084UL6 oryza sativ	229	33	62.3	191	10	08W2T0	08W2T0 oryza sativ
157	34	64.2	1030	10	07XR31	07XR31 oryza sativ	230	33	62.3	191	10	07XFA4	07XFA4 oryza sativ
158	34	64.2	1031	10	07X913	07X913 oryza sativ	231	33	62.3	193	5	077233	077233 trypanosoma
159	34	64.2	1032	10	07XSX8	07XSX8 oryza sativ	232	33	62.3	201	4	09HBL4	09HBL4 homo sapien
160	34	64.2	1050	10	07XNN8	07XNN8 oryza sativ	233	33	62.3	206	5	08MUE7	08MUE7 chlamys far
161	34	64.2	1052	10	07X6B2	07X6B2 oryza sativ	234	33	62.3	211	15	08Q656	08Q656 human immun
162	34	64.2	1063	10	08LMM2	08LMM2 oryza sativ	235	33	62.3	213	16	09Z7M6	09Z7M6 chlamydia p

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:01:12 ; Search time 40 Seconds  
(without alignments)  
78.880 Million cell updates/sec

Title: US-10-697-055-7  
Perfect score: 53  
Sequence: 1 TPRVTGGGAM 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : SPTREMBL\_25.\*

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_plant:\*  
10: sp\_rodent:\*  
11: sp\_virus:\*  
12: sp\_virid:\*  
13: sp\_virid:\*  
14: sp\_virid:\*  
15: sp\_virid:\*  
16: sp\_bacterioph:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	561	12	Q7ter5 human cytom
2	39	73.6	255	11	Q8c1l4 mus musculu
3	39	73.6	339	11	Q9l44 mus musculu
4	38	71.7	340	11	Q8JUV6
5	38	71.7	383	10	Q9SYG2
6	37	69.8	258	16	Q7wigo bordetella
7	37	69.8	258	16	Q7w6i8 bordetella
8	37	69.8	258	16	Q7vym1 bordetella
9	37	69.8	264	16	Q7v426 prochloroco
10	37	69.8	345	16	Q7V426
11	37	69.8	379	10	Q9FE35
12	37	69.8	379	10	Q7XAA2
13	37	69.8	653	16	Q9ABH5
14	37	69.8	748	16	Q82D09
15	36	67.9	54	16	Q82C77
16	36	67.9	85	16	Q92ZW2

17	36	67.9	127	2	Q46488	Q46488 leifsonia x
18	36	67.9	139	10	Q84QG3	Q84QG3 zea mays su
19	36	67.9	139	10	Q84QG2	Q84QG2 zea mays su
20	36	67.9	139	10	Q84QG1	Q84QG1 zea mays su
21	36	67.9	139	10	Q84P41	Q84P41 zea mays (s
22	36	67.9	139	10	Q84P40	Q84P40 zea mays (s
23	36	67.9	139	10	Q84P39	Q84P39 zea mays (s
24	36	67.9	139	10	Q84P38	Q84P38 zea mays (s
25	36	67.9	139	10	Q84P37	Q84P37 zea mays (s
26	36	67.9	139	10	Q84P36	Q84P36 zea mays (s
27	36	67.9	139	10	Q84J06	Q84J06 zea mays su
28	36	67.9	181	16	Q89GB3	Q89GB3 bradyrhizob
29	36	67.9	222	2	Q8KVT4	Q8KVT4 streptomyc
30	36	67.9	244	16	Q825D0	Q825D0 streptomyc
31	36	67.9	261	5	Q9W3C6	Q9W3C6 drosophila
32	36	67.9	262	9	Q9T178	Q9T178 bacterioph
33	36	67.9	354	2	Q9R1P7	Q9R1P7 streptomyc
34	36	67.9	361	16	Q8G501	Q8G501 bifidobacte
35	36	67.9	383	10	Q9ARP7	Q9ARP7 oryza sativ
36	36	67.9	388	16	Q69817	Q69817 streptomyc
37	36	67.9	501	12	Q8B329	Q8B329 sweet potat
38	36	67.9	501	12	Q8B328	Q8B328 sweet potat
39	36	67.9	501	12	Q8B327	Q8B327 sweet potat
40	36	67.9	501	12	Q8B326	Q8B326 sweet potat
41	36	67.9	501	12	Q8B325	Q8B325 sweet potat
42	36	67.9	501	12	Q8B323	Q8B323 sweet potat
43	36	67.9	788	16	Q67557	Q67557 aquifex aeo
44	36	67.9	799	10	Q81387	Q81387 zea mays (m
45	36	67.9	808	2	Q8G8W9	Q8G8W9 pseudomonas
46	36	67.9	823	2	Q8GN89	Q8GN89 pseudomonas
47	36	67.9	879	16	Q9JUL2	Q9JUL2 bradyrhizob
48	36	67.9	1540	4	O15080	O15080 homo sapien
49	36	67.9	3456	12	P83201	P83201 sweet potat
50	35	66.0	126	11	Q9CRJ6	Q9CRJ6 mus musculu
51	35	66.0	127	4	Q96FF4	Q96FF4 homo sapien
52	35	66.0	228	10	Q8H543	Q8H543 oryza sativ
53	35	66.0	233	10	Q8W0F2	Q8W0F2 oryza sativ
54	35	66.0	280	16	Q8PJ96	Q8PJ96 xanthomonas
55	35	66.0	290	10	Q9AXM0	Q9AXM0 brassica na
56	35	66.0	304	16	Q8UCB3	Q8UCB3 agrobacteri
57	35	66.0	330	16	Q82MF0	Q82MF0 streptomyc
58	35	66.0	336	16	Q9X1N0	Q9X1N0 thermotoga
59	35	66.0	359	11	Q9JMD5	Q9JMD5 mus musculu
60	35	66.0	359	11	Q9JH8	Q9JH8 mus musculu
61	35	66.0	359	11	Q9JIE7	Q9JIE7 mus musculu
62	35	66.0	359	11	Q8C1I9	Q8C1I9 mus musculu
63	35	66.0	361	3	Q9C2H7	Q9C2H7 neurospora
64	35	66.0	366	4	Q8NF40	Q8NF40 homo sapien
65	35	66.0	370	11	Q9J1Q7	Q9J1Q7 mus musculu
66	35	66.0	424	16	Q87264	Q87264 pseudomonas
67	35	66.0	432	2	Q8GGR5	Q8GGR5 streptomyc
68	35	66.0	444	16	Q82QY2	Q82QY2 streptomyc
69	35	66.0	458	5	Q8MQ76	Q8MQ76 caenorhabdi
70	35	66.0	475	5	Q8MPU3	Q8MPU3 caenorhabdi
71	35	66.0	478	10	Q8S277	Q8S277 oryza sativ
72	35	66.0	485	5	Q9GVA4	Q9GVA4 styela clav
73	35	66.0	489	5	Q9GQ00	Q9GQ00 caenorhabdi
74	35	66.0	522	5	Q9XTB0	Q9XTB0 caenorhabdi
75	35	66.0	522	5	Q9BH43	Q9BH43 caenorhabdi
76	35	66.0	539	3	Q03761	Q03761 saccharomyc
77	35	66.0	586	10	Q9FJ99	Q9FJ99 arabidopsis
78	35	66.0	718	5	Q9NKS6	Q9NKS6 leishmania
79	35	66.0	923	16	O53890	O53890 mycobacteri
80	35	66.0	923	16	Q7U0X8	Q7U0X8 mycobacteri
81	35	66.0	1127	5	Q7T1U1	Q7T1U1 caenorhabdi
82	35	66.0	1138	16	Q92KB6	Q92KB6 rhizobium m
83	35	66.0	1161	5	Q8WQA0	Q8WQA0 caenorhabdi
84	35	66.0	1237	16	Q7UFY1	Q7UFY1 rhodospirill
85	35	66.0	1257	16	Q8CVB2	Q8CVB2 oceanobacil
86	35	66.0	1433	16	Q8DLX7	Q8DLX7 synechococc
87	35	66.0	3146	5	Q9VUB5	Q9VUB5 drosophila
88	35	66.0	3643	16	Q9RK14	Q9RK14 streptomyc
89	34	64.2	46	12	O92573	O92573 hepatitis c

APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4695  
LENGTH: 263  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4695

Query Match 64.2%; Score 34; DB 9; Length 263;  
Best Local Similarity 75.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RVTGGGAM 10  
Db 86 RMSGGGAM 93

RESULT 50  
US-10-102-806-730  
Sequence 730, Application US/10102806  
Publication No. US2003005421A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103P1C1  
CURRENT APPLICATION NUMBER: US/10/102,806  
CURRENT FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: 09/925,298  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 730  
LENGTH: 288  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-102-806-730

Query Match 64.2%; Score 34; DB 14; Length 288;  
Best Local Similarity 66.7%; Pred. No. 5.2e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
Db 10 PRAPGGGAM 18

Search completed: March 7, 2004, 13:09:18  
Job time : 43 secs

```
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 262
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-262

Query Match 64.2%; Score 34; DB 14; Length 128;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
DB 9 PKVPGGGA 16

RESULT 46
US-10-369-493-17800
; Sequence 17800, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17800
; LENGTH: 148
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17800

Query Match 64.2%; Score 34; DB 15; Length 148;
Best Local Similarity 72.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 TPRV--TGGGA 9
DB 70 TPKVIATGGGA 80

RESULT 47
US-10-247-671-173
; Sequence 173, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
```

```
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 173
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 3282941CD1
US-10-247-671-173

Query Match 64.2%; Score 34; DB 14; Length 206;
Best Local Similarity 85.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
DB 75 PRVAGGG 81

RESULT 48
US-09-738-626-4633
; Sequence 4633, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4633
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4633

Query Match 64.2%; Score 34; DB 9; Length 260;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 2 PRVTGGG 8
DB 88 PRISGGG 94

RESULT 49
US-09-738-626-4695
; Sequence 4695, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
```

```
; SEQ ID NO 6240
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6240

Query Match      66.0%; Score 35; DB 15; Length 522;
Best Local Similarity 70.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
Db 214 TGRMRGGGAM 223

RESULT 41
US-10-369-493-6241
; Sequence 6241, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6241
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6241

Query Match      66.0%; Score 35; DB 15; Length 522;
Best Local Similarity 70.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
Db 214 TGRMRGGGAM 223

RESULT 42
US-10-227-035-2
; Sequence 2, Application US/10227035
; Publication No. US20030082809A1
; GENERAL INFORMATION:
; APPLICANT: Quail, Peter H.
; APPLICANT: Huq, Enamul
; APPLICANT: Tepperman, James
; APPLICANT: Sato, Sae
; TITLE OF INVENTION: A UNIVERSAL LIGHT-SWITCHABLE GENE
; TITLE OF INVENTION: PROMOTER SYSTEM
; FILE REFERENCE: 416272061500
; CURRENT APPLICATION NUMBER: US/10/227,035
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/355,402
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/314,615
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-227-035-2

Query Match      66.0%; Score 35; DB 14; Length 524;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGG 8
Db 470 PRVSGGG 476

RESULT 43
US-10-219-220-71
; Sequence 71, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-71

Query Match      64.2%; Score 34; DB 14; Length 86;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
Db 9 PKVPGGGA 16

RESULT 44
US-10-108-260A-2675
; Sequence 2675, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2675
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2675

Query Match      64.2%; Score 34; DB 15; Length 111;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGG 7
Db 94 TPRATGG 100

RESULT 45
US-10-219-220-262
; Sequence 262, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
```



; APPLICANT: Wei, Zhong-Min  
; APPLICANT: Schading, Richard L.  
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR-INDUCED STRESS  
; TITLE OF INVENTION: RESISTANCE  
; FILE REFERENCE: 21829/203 (EBC-003)  
; CURRENT APPLICATION NUMBER: US/10/441,736  
; CURRENT FILING DATE: 2003-05-20  
; PRIOR APPLICATION NUMBER: 60/107,243  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 09/431,614  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas syringae  
US-10-441-736-14

Query Match 66.0%; Score 35; DB 15; Length 424;  
Best Local Similarity 66.7%; Pred. No. 5.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
|||  
Db 169 TPTATGGGS 177

RESULT 37  
US-10-314-657-23  
; Sequence 23, Application US/10314657  
; Publication No. US2003017588A1  
; GENERAL INFORMATION:  
; APPLICANT: SHEN, Ben  
; APPLICANT: CHENG, Yi-Qiang  
; APPLICANT: TANG, Gong-Li  
; TITLE OF INVENTION: Synchro Acyltransferases Associated with Type I Polyketide  
; TITLE OF INVENTION: Synthase Acyltransferases and Methods of Use  
; FILE REFERENCE: 054030-0021  
; CURRENT APPLICATION NUMBER: US/10/314,657  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: PCT/US02/08937  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 60/278,935  
; PRIOR FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 214  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 23  
; LENGTH: 432  
; TYPE: PRT  
; ORGANISM: Streptomyces atroolivaceus  
US-10-314-657-23

Query Match 66.0%; Score 35; DB 14; Length 432;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
|  
Db 129 PDVTGGGSL 137

RESULT 38  
US-10-156-761-7910  
; Sequence 7910, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 7910  
; LENGTH: 444  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-7910

Query Match 66.0%; Score 35; DB 14; Length 444;  
Best Local Similarity 66.7%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 2 PRVTGGGAM 10  
|||  
Db 373 PRAAGGGVM 381

RESULT 39  
US-10-032-585-7677  
; Sequence 7677, Application US/10032585  
; Publication No. US20030180953A1  
; GENERAL INFORMATION:  
; APPLICANT: Terry, Roemer D.  
; APPLICANT: Bo, Jiang  
; APPLICANT: Charles, Boone  
; APPLICANT: Howard, Bussey  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
; FILE REFERENCE: 10182-005-999  
; CURRENT APPLICATION NUMBER: US/10/032,585  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 8000  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7677  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-10-032-585-7677

Query Match 66.0%; Score 35; DB 14; Length 502;  
Best Local Similarity 75.0%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8  
|||  
Db 60 TPKVTGSG 67

RESULT 40  
US-10-369-493-6240  
; Sequence 6240, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374

```
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-895-593-2

Query Match      66.0%; Score 35; DB 9; Length 370;
Best Local Similarity 77.8%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TPRVTGGGA 10
Db 344 TPTATGGGS 352

RESULT 32
US-09-835-684-9
; Sequence 9, Application US/09835684
; Patent No. US20020019337A1
; GENERAL INFORMATION:
; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Remick, Dean
; TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
; TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
; TITLE OF INVENTION: DESICCATION
; FILE REFERENCE: 21829/71
; CURRENT APPLICATION NUMBER: US/09/835,684
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/198,359
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-835-684-9

Query Match      66.0%; Score 35; DB 9; Length 424;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 169 TPTATGGGS 177

RESULT 33
US-09-880-371-9
; Sequence 9, Application US/09880371
; Patent No. US20020059658A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Derocher, Jay
; TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 21829/91
; CURRENT APPLICATION NUMBER: US/09/880,371
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/211,585
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-880-371-9

Query Match      66.0%; Score 35; DB 9; Length 424;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 169 TPTATGGGS 177

RESULT 34
US-09-879-248-14
; Sequence 14, Application US/09879248
; Patent No. US20020062500A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 21829/81
; CURRENT APPLICATION NUMBER: US/09/879,248
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/212,211
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-879-248-14

Query Match      66.0%; Score 35; DB 9; Length 424;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 169 TPTATGGGS 177

RESULT 35
US-10-010-390-9
; Sequence 9, Application US/10010390
; Publication No. US20030104979A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Leon, Ernesto
; APPLICANT: Oviedo, Agustín
; TITLE OF INVENTION: METHODS OF INHIBITING DESICCATION OF CUTTINGS REMOVED
; TITLE OF INVENTION: FROM ORNAMENTAL PLANTS
; FILE REFERENCE: 21829/111
; CURRENT APPLICATION NUMBER: US/10/010,390
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/248,169
; PRIOR FILING DATE: 2000-11-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-010-390-9

Query Match      66.0%; Score 35; DB 14; Length 424;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPRVTGGGA 9
Db 169 TPTATGGGS 177

RESULT 36
US-10-441-736-14
; Sequence 14, Application US/10441736
; Publication No. US20040016029A1
; GENERAL INFORMATION:
```

; PRIOR FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: TRANSMEM  
; LOCATION: (227)..(247)  
US-09-895-943-3

Query Match 66.0%; Score 35; DB 9; Length 353;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
||| |||||  
Db 327 PRGPGGGAM 335

## RESULT 28

US-09-895-593-3  
; Sequence 3, Application US/09895593  
; Patent No. US20020160949A1  
; GENERAL INFORMATION:  
; APPLICANT: Pandey, Akhilesh  
; APPLICANT: Ozaki, Katsutoshi  
; APPLICANT: Baumann, Heinz  
; APPLICANT: Levin, Steven D.  
; APPLICANT: Farr, Andrew G.  
; APPLICANT: Ziegler, Steven F.  
; APPLICANT: Leonard, Warren J.  
; APPLICANT: Lodish, Harvey F.  
; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and  
; FILE REFERENCE: 00-514-E  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/215,658  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: TRANSMEM  
; LOCATION: (227)..(247)  
US-09-895-593-3

Query Match 66.0%; Score 35; DB 9; Length 353;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
||| |||||  
Db 327 PRGPGGGAM 335

## RESULT 29

US-10-226-872-7  
; Sequence 7, Application US/10226872  
; Publication No. US20030157118A1  
; GENERAL INFORMATION:  
; APPLICANT: Cabezón-Silva, Teresa Elisa Virginia  
; APPLICANT: Coche, Thierry  
; APPLICANT: Gaulis, Swann Romain Jean-Thomas  
; APPLICANT: Vinals Y De Bassols, Carlota  
; APPLICANT: Cassart, Jean-Pol  
; TITLE OF INVENTION: Tumour-Specific Animal Proteins  
; FILE REFERENCE: BC45300-1

; CURRENT APPLICATION NUMBER: US/10/226,872  
; CURRENT FILING DATE: 2002-12-05  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Human  
US-10-226-872-7

Query Match 66.0%; Score 35; DB 14; Length 361;  
Best Local Similarity 75.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9  
||| |||||  
Db 79 PRLAGGA 86

## RESULT 30

US-09-895-943-2  
; Sequence 2, Application US/09895943  
; Patent No. US20020068323A1  
; GENERAL INFORMATION:  
; APPLICANT: Saris, Chris  
; APPLICANT: Chang, Ming-Shi  
; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and  
; FILE REFERENCE: 00-514-C  
; CURRENT APPLICATION NUMBER: US/09/895,943  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/214,866  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-895-943-2

Query Match 66.0%; Score 35; DB 9; Length 370;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRVTGGGAM 10  
||| |||||  
Db 344 PRGPGGGAM 352

## RESULT 31

US-09-895-593-2  
; Sequence 2, Application US/09895593  
; Patent No. US20020160949A1  
; GENERAL INFORMATION:  
; APPLICANT: Pandey, Akhilesh  
; APPLICANT: Ozaki, Katsutoshi  
; APPLICANT: Baumann, Heinz  
; APPLICANT: Levin, Steven D.  
; APPLICANT: Farr, Andrew G.  
; APPLICANT: Ziegler, Steven F.  
; APPLICANT: Leonard, Warren J.  
; APPLICANT: Lodish, Harvey F.  
; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and  
; FILE REFERENCE: 00-514-E  
; CURRENT APPLICATION NUMBER: US/09/895,593  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/215,658  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2

DB 22 VTGGGAM 28

US-10-374-780A-1701

Sequence 1701, Application US/10374780A

Publication No. US20040019927A1

GENERAL INFORMATION:

APPLICANT: Sherman, Bradley K

APPLICANT: Riechmann, Jose Luis

APPLICANT: Jiang, Cai-Zhong

APPLICANT: Heard, Jacqueline E

APPLICANT: Haake, Volker

APPLICANT: Creelman, Robert A

APPLICANT: Ratcliffe, Oliver

APPLICANT: Adam, Luc J

APPLICANT: Reuber, T. Lynne

APPLICANT: Keddle, James

APPLICANT: Brown, Pierre E

APPLICANT: Pilgrim, Marsha L

APPLICANT: Dubell III, Arnold T

APPLICANT: Pineda, Omaira

APPLICANT: Yu, Guo-Liang

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

FILE REFERENCE: MBI-0047 CIP

CURRENT APPLICATION NUMBER: US/10/374,780A

CURRENT FILING DATE: 2003-02-25

PRIOR APPLICATION NUMBER: 09/837,944

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/310,847

PRIOR FILING DATE: 2001-08-09

PRIOR APPLICATION NUMBER: 09/934,455

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/336,049

PRIOR FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/338,692

PRIOR FILING DATE: 2001-12-11

PRIOR APPLICATION NUMBER: 10/171,468

PRIOR FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 10/225,066

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,067

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,068

PRIOR FILING DATE: 2002-08-09

NUMBER OF SEQ ID NOS: 2906

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1701

LENGTH: 338

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Orthologous to G1652

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8

DB 91 SPHVITGGG 98

RESULT 27

US-09-895-943-3

Sequence 3, Application US/09895943

Patent No. US20020068323A1

GENERAL INFORMATION:

APPLICANT: Saris, Chris

APPLICANT: Chang, Ming-Shi

TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and

TITLE OF INVENTION: Uses Thereof

FILE REFERENCE: 00-514-C

CURRENT APPLICATION NUMBER: US/09/895,943

CURRENT FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,866

US-10-374-780A-1701

Query Match 66.0%; Score 35; DB 15; Length 338;

Best Local Similarity 75.0%; Pred. No. 4.1e+02

Query Match 67.9%; Score 36; DB 15; Length 276;  
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
 |||:|:|:  
 DB 228 TPRLTGSGS 236

RESULT 22

US-10-369-493-15324  
 ; Sequence 15324, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 15324  
 ; LENGTH: 294  
 ; TYPE: PRT  
 ; ORGANISM: Xanthomonas campestris  
 ; US-10-369-493-15324

Query Match 67.9%; Score 36; DB 15; Length 294;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
 |||:|:|:  
 DB 242 TPRLTGSGS 250

RESULT 23

US-10-369-493-20821  
 ; Sequence 20821, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 20821  
 ; LENGTH: 600  
 ; TYPE: PRT  
 ; ORGANISM: Rhodospseudomonas palustris  
 ; US-10-369-493-20821

Query Match 67.9%; Score 36; DB 15; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VTGGGAM 10  
 |||:|:|:  
 DB 228 TPRLTGSGS 236

Query Match 67.9%; Score 36; DB 14; Length 244;  
 Best Local Similarity 70.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10  
 |||:|:|:  
 DB 74 TPRPTGGAKM 83

RESULT 20

US-10-369-493-15690  
 ; Sequence 15690, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 15690  
 ; LENGTH: 276  
 ; TYPE: PRT  
 ; ORGANISM: Xanthomonas campestris  
 ; US-10-369-493-15690

Query Match 67.9%; Score 36; DB 15; Length 276;  
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9  
 |||:|:|:  
 DB 228 TPRLTGSGS 236

RESULT 21

US-10-369-493-16082  
 ; Sequence 16082, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 16082  
 ; LENGTH: 276  
 ; TYPE: PRT  
 ; ORGANISM: Xanthomonas campestris  
 ; US-10-369-493-16082

```

; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-306-762-109

Query Match
Best Local Similarity 59.8%; Score 37; DB 14; Length 313;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
Db 98 SPKVTGGG 105

RESULT 16
US-10-156-761-12713
; Sequence 12713, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12713
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12713

Query Match
Best Local Similarity 69.8%; Score 37; DB 14; Length 748;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
Db 593 PRINGGGA 600

RESULT 17
US-10-156-761-13012
; Sequence 13012, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12713
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12713

Query Match
Best Local Similarity 69.8%; Score 37; DB 14; Length 748;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
Db 593 PRINGGGA 600

RESULT 18
US-10-029-386-27829
; Sequence 27829, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27829
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133476.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
; OTHER INFORMATION: SWISSPROT HIT: Q08788, EVALUATE 1.60e+00
US-10-029-386-27829

Query Match
Best Local Similarity 67.9%; Score 36; DB 14; Length 145;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
Db 25 TPRATGRGA 33

RESULT 19
US-10-156-761-15056
; Sequence 15056, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12713
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12713

Query Match
Best Local Similarity 67.9%; Score 36; DB 14; Length 145;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
Db 25 TPRATGRGA 33

```

us-10-697-055-7.rapb

Sun Mar 7 13:59:03 2004

```
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-051-874-87

Query Match 73.6%; Score 39; DB 15; Length 342;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9
DB 43 PRITGGGS 50

RESULT 13
US-10-259-194A-124
; Sequence 124, Application US/10259194A
; Publication No. US20040010815A1
; GENERAL INFORMATION:
; APPLICANT: Lange, Markus B.
; APPLICANT: Ghassemtan, Majid
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF PLANT GENES
; FILE REFERENCES: 70029-NP
; CURRENT APPLICATION NUMBER: US/10/259,194A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,743
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 662
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 124
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-259-194A-124

Query Match 73.6%; Score 39; DB 15; Length 665;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
DB 27 TPRLSGGGTL 36

RESULT 14
US-10-374-780A-2854
; Sequence 2854, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pinada, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCES: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2854
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2373
US-10-374-780A-2854

Query Match 71.7%; Score 38; DB 15; Length 383;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPRVTGGG 8
DB 62 TPSVTGGG 69

RESULT 15
US-10-306-762-109
; Sequence 109, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; FILE REFERENCES: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCES: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
```

```

; SEQ ID NO 86
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-051-874-86

Query Match      73.6%; Score 39; DB 15; Length 342;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 PRVTGGGA 9
Db      43 PRITGGGS 50

RESULT 12
US-10-051-874-86
; Sequence 86, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sash
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zernhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 86
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-051-874-86

Query Match      73.6%; Score 39; DB 15; Length 342;
Best Local Similarity 75.0%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 PRVTGGGA 9
Db      43 PRITGGGS 50

RESULT 12
US-10-051-874-86
; Sequence 87, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sash
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zernhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1

```



;; CURRENT FILING DATE: 2003-04-03  
;; PRIOR APPLICATION NUMBER: US/09/692,170C  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: US 09/534,639  
;; PRIOR FILING DATE: 2000-03-27  
;; PRIOR APPLICATION NUMBER: US 09/075,257  
;; PRIOR FILING DATE: 1998-05-11  
;; PRIOR APPLICATION NUMBER: US 09/021,298  
;; PRIOR FILING DATE: 1998-02-10  
;; PRIOR APPLICATION NUMBER: US 08/950,064  
;; PRIOR FILING DATE: 1997-10-14  
;; PRIOR APPLICATION NUMBER: US 08/747,488  
;; PRIOR FILING DATE: 1996-11-12  
;; NUMBER OF SEQ ID NOS: 44  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 8  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Artificial sequence  
;; OTHER INFORMATION: Variant human cytomegalovirus peptide epitope  
;; FEATURE:  
;; NAME/KEY: MISC FEATURE  
;; LOCATION: (10)..(10)  
;; OTHER INFORMATION: xaa = L, F or M  
US-10-405-231-8

Query Match 90.6%; Score 48; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRTVGGGA 9  
|||||||  
DB 1 TPRTVGGGA 9

## RESULT 9

US-10-051-874-89  
;; Sequence 89, Application US/10051874  
;; Publication No. US20040005557A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Padigaru, Muralidhara  
;; APPLICANT: Alsobrook II, John P  
;; APPLICANT: Colman, Steven D  
;; APPLICANT: Spytek, Kimberly A  
;; APPLICANT: Boldog, Ferenc  
;; APPLICANT: Vernet, Corine AM  
;; APPLICANT: Li, Li  
;; APPLICANT: Shenoy, Suresh G  
;; APPLICANT: Casman, Stacie J  
;; APPLICANT: Guo, Xiaojia Sasha  
;; APPLICANT: Edinger, Shlomit R  
;; APPLICANT: MacDougall, John R  
;; APPLICANT: Malyankar, Uriel M  
;; APPLICANT: Patturajan, Meera  
;; APPLICANT: Shinkets, Richard A  
;; APPLICANT: Pena, Carol EA  
;; APPLICANT: Tchernev, Velizar T  
;; APPLICANT: Zerhusen, Bryan D  
;; APPLICANT: Millet, Isabelle  
;; APPLICANT: Miller, Charles E  
;; APPLICANT: Lepley, Denise M  
;; APPLICANT: Smithson, Glenna  
;; APPLICANT: Baumgartner, Jason C  
;; APPLICANT: Herrman, John L  
;; APPLICANT: Feyman, John A  
;; APPLICANT: Gorman, Linda  
;; APPLICANT: Mezes, Peter D  
;; APPLICANT: Kekuda, Ramesh  
;; APPLICANT: Taupier Jr, Raymond J  
;; APPLICANT: Gerlach, Valerie  
;; APPLICANT: Grosse, William M  
;; APPLICANT: Liu, Xiachong

;; APPLICANT: Ellerman, Karen  
;; APPLICANT: Rothenberg, Mark  
;; APPLICANT: Stone, David J  
;; APPLICANT: Burgess, Catherine E  
;; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
;; FILE REFERENCE: 21402-245  
;; CURRENT APPLICATION NUMBER: US/10/051,874  
;; CURRENT FILING DATE: 2002-09-25  
;; PRIOR APPLICATION NUMBER: 60/268,595  
;; PRIOR FILING DATE: 2001-02-14  
;; PRIOR APPLICATION NUMBER: 60/325,306  
;; PRIOR FILING DATE: 2001-09-27  
;; PRIOR APPLICATION NUMBER: 60/262,587  
;; PRIOR FILING DATE: 2001-01-18  
;; PRIOR APPLICATION NUMBER: 60/272,409  
;; PRIOR FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: 60/262,454  
;; PRIOR FILING DATE: 2001-01-18  
;; PRIOR APPLICATION NUMBER: 60/276,777  
;; PRIOR FILING DATE: 2001-03-16  
;; PRIOR APPLICATION NUMBER: 60/291,672  
;; PRIOR FILING DATE: 2001-05-17  
;; PRIOR APPLICATION NUMBER: 60/330,336  
;; PRIOR FILING DATE: 2001-10-18  
;; PRIOR APPLICATION NUMBER: 60/265,530  
;; PRIOR FILING DATE: 2001-01-31  
;; PRIOR APPLICATION NUMBER: 60/261,376  
;; PRIOR FILING DATE: 2001-01-16  
;; NUMBER OF SEQ ID NOS: 269  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 89  
;; LENGTH: 285  
;; TYPE: PRT  
;; ORGANISM: Mus musculus  
US-10-051-874-89

Query Match 73.6%; Score 39; DB 15; Length 285;  
Best Local Similarity 75.0%; Pred. No. 75;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRVTGGGA 9  
|||||  
DB 43 PRITGGGS 50

## RESULT 10

US-10-109-616-2  
;; Sequence 2, Application US/10109616  
;; Publication No. US20030167484A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Allen, Keith D.  
;; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING CHANNEL  
;; TITLE OF INVENTION: ACTIVATING PROTEASE 1 (CAP1) GENE DISRUPTIONS  
;; FILE REFERENCE: R-490  
;; CURRENT APPLICATION NUMBER: US/10/109,616  
;; CURRENT FILING DATE: 2002-03-28  
;; PRIOR APPLICATION NUMBER: US 60/280,509  
;; PRIOR FILING DATE: 2001-03-29  
;; PRIOR APPLICATION NUMBER: US 60/311,055  
;; PRIOR FILING DATE: 2001-08-08  
;; NUMBER OF SEQ ID NOS: 4  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 2  
;; LENGTH: 339  
;; TYPE: PRT  
;; ORGANISM: Mus musculus  
US-10-109-616-2

Query Match 73.6%; Score 39; DB 14; Length 339;  
Best Local Similarity 75.0%; Pred. No. 89;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/223,538
; FILING DATE: 19-AUG-2002
; CLASSIFICATION: <unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-223-538-6
Query Match 100.0%; Score 53; DB 14; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
DB 407 TPRVTGGGAM 416

RESULT 6
US-10-223-538-8
; Sequence 8, Application US/10223538
; Publication No. US20030120060A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsei, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. US20030120060A1el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/223,538
; FILING DATE: 19-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
```

```
;
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-223-538-8
Query Match 100.0%; Score 53; DB 14; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGAM 10
DB 407 TPRVTGGGAM 416

RESULT 7
US-10-238-607-8
; Sequence 8, Application US/10238607
; Publication No. US20030118602A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: HCMV- REACTIVE T CELLS AND USES THEREFOR
; FILE REFERENCE: 1954-398
; CURRENT APPLICATION NUMBER: US/10/238,607
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 09/692,170
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,639
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/075,257
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: US 09/021,298
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 08/950,064
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: US 08/747,488
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Variant human cytomegalovirus peptide epitope
; NAME/KEY: MISC FEATURE
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa = L, F or M
US-10-238-607-8
Query Match 90.6%; Score 48; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPRVTGGGA 9
DB 1 TPRVTGGGA 9

RESULT 8
US-10-405-231-8
; Sequence 8, Application US/10405231
; Publication No. US20030190328A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-346
; CURRENT APPLICATION NUMBER: US/10/405,231
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-079A-2

Query Match      100.0%; Score 53; DB 9; Length 561;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGAM 10
        |||||
Db      417 TPRVTGGGAM 426

RESULT 4
US-10-434-982-2
; Sequence 2, Application US/10434982
; Publication No. US20030199673A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS COMPOUNDS
; FILE REFERENCE: 5018C
; CURRENT APPLICATION NUMBER: US/10/434,982
; CURRENT FILING DATE: 2003-05-10
; PRIOR APPLICATION NUMBER: US 09/812,079
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: US 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-434-982-2

Query Match      100.0%; Score 53; DB 14; Length 561;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGAM 10
        |||||
Db      417 TPRVTGGGAM 426

RESULT 5
US-10-223-538-6
; Sequence 6, Application US/10223538
; Publication No. US20030120060A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. US20030120060A1 Cytomegalovirus DNA Constructs and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

---

```
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human cytomegalovirus
US-10-238-607-7

Query Match      100.0%; Score 53; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGAM 10
        |||||
Db      1 TPRVTGGGAM 10

RESULT 2
US-10-405-231-7
; Sequence 7, Application US/10405231
; Publication No. US20030190328A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS
; FILE REFERENCE: 1954-346
; CURRENT APPLICATION NUMBER: US/10/405,231
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US/09/692,170C
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,639
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/075,257
; PRIOR FILING DATE: 1998-05-11
; PRIOR APPLICATION NUMBER: US 09/021,298
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 08/950,064
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: US 08/747,488
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human cytomegalovirus
US-10-405-231-7

Query Match      100.0%; Score 53; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TPRVTGGGAM 10
        |||||
Db      1 TPRVTGGGAM 10

RESULT 3
US-09-812-079A-2
; Sequence 2, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
```

381	34	64.2	346	14	US-10-147-490-432	Sequence 432, App
382	34	64.2	346	14	US-10-147-494-432	Sequence 432, App
383	34	64.2	346	14	US-10-147-498-432	Sequence 432, App
384	34	64.2	346	14	US-10-147-514-432	Sequence 432, App
385	34	64.2	346	14	US-10-147-524-432	Sequence 432, App
386	34	64.2	346	14	US-10-152-379-432	Sequence 432, App
387	34	64.2	346	14	US-10-152-394-432	Sequence 432, App
388	34	64.2	346	14	US-10-152-406-432	Sequence 432, App
389	34	64.2	346	14	US-10-156-847-432	Sequence 432, App
390	34	64.2	346	14	US-10-157-778-432	Sequence 432, App
391	34	64.2	346	14	US-10-157-799-432	Sequence 432, App
392	34	64.2	346	14	US-10-160-504-432	Sequence 432, App
393	34	64.2	346	14	US-10-145-634-432	Sequence 432, App
394	34	64.2	346	14	US-10-147-520-432	Sequence 432, App
395	34	64.2	346	14	US-10-157-781-432	Sequence 432, App
396	34	64.2	346	14	US-10-176-989-432	Sequence 432, App
397	34	64.2	346	14	US-10-147-491-432	Sequence 432, App
398	34	64.2	346	14	US-10-152-378-432	Sequence 432, App
399	34	64.2	346	14	US-10-152-382-432	Sequence 432, App
400	34	64.2	346	14	US-10-153-756-432	Sequence 432, App
401	34	64.2	346	14	US-10-152-383-432	Sequence 432, App
402	34	64.2	346	14	US-10-152-384-432	Sequence 432, App
403	34	64.2	346	14	US-10-152-387-432	Sequence 432, App
404	34	64.2	346	14	US-10-152-389-432	Sequence 432, App
405	34	64.2	346	14	US-10-152-390-432	Sequence 432, App
406	34	64.2	346	14	US-10-152-392-432	Sequence 432, App
407	34	64.2	346	14	US-10-153-756-432	Sequence 432, App
408	34	64.2	346	14	US-10-157-797-432	Sequence 432, App
409	34	64.2	346	14	US-10-158-491-432	Sequence 432, App
410	34	64.2	346	14	US-10-142-762-432	Sequence 432, App
411	34	64.2	346	14	US-10-142-764-432	Sequence 432, App
412	34	64.2	346	14	US-10-145-625-432	Sequence 432, App
413	34	64.2	346	14	US-10-145-627-432	Sequence 432, App
414	34	64.2	346	14	US-10-145-960-432	Sequence 432, App
415	34	64.2	346	14	US-10-145-962-432	Sequence 432, App
416	34	64.2	346	14	US-10-146-789-432	Sequence 432, App
417	34	64.2	346	14	US-10-147-496-432	Sequence 432, App
418	34	64.2	346	14	US-10-147-498-432	Sequence 432, App
419	34	64.2	346	14	US-10-147-505-432	Sequence 432, App
420	34	64.2	346	14	US-10-147-516-432	Sequence 432, App
421	34	64.2	346	14	US-10-152-398-432	Sequence 432, App
422	34	64.2	346	14	US-10-139-980-432	Sequence 432, App
423	34	64.2	346	14	US-10-145-750-432	Sequence 432, App
424	34	64.2	346	14	US-10-152-373-432	Sequence 432, App
425	34	64.2	346	14	US-10-223-081-130	Sequence 430, App
426	34	64.2	346	14	US-10-121-044-432	Sequence 432, App
427	34	64.2	346	14	US-10-121-055-432	Sequence 432, App
428	34	64.2	346	14	US-10-121-057-432	Sequence 432, App
429	34	64.2	346	14	US-10-121-058-432	Sequence 432, App
430	34	64.2	346	14	US-10-121-059-432	Sequence 432, App
431	34	64.2	346	14	US-10-121-060-432	Sequence 432, App
432	34	64.2	346	14	US-10-123-109-432	Sequence 432, App
433	34	64.2	346	14	US-10-123-154-432	Sequence 432, App
434	34	64.2	346	14	US-10-123-157-432	Sequence 432, App
435	34	64.2	346	14	US-10-123-906-432	Sequence 432, App
436	34	64.2	346	14	US-10-124-814-432	Sequence 432, App
437	34	64.2	346	14	US-10-124-816-432	Sequence 432, App
438	34	64.2	346	14	US-10-124-820-432	Sequence 432, App
439	34	64.2	346	14	US-10-125-704-432	Sequence 432, App
440	34	64.2	346	14	US-10-125-927-432	Sequence 432, App
441	34	64.2	346	14	US-10-223-082-130	Sequence 430, App
442	34	64.2	346	14	US-10-142-889-432	Sequence 432, App
443	34	64.2	346	14	US-10-145-874-432	Sequence 432, App
444	34	64.2	346	14	US-10-147-497-432	Sequence 432, App
445	34	64.2	346	14	US-10-152-371-432	Sequence 432, App
446	34	64.2	346	14	US-10-152-374-432	Sequence 432, App
447	34	64.2	346	14	US-10-152-375-432	Sequence 432, App
448	34	64.2	346	14	US-10-152-377-432	Sequence 432, App
449	34	64.2	346	14	US-10-152-386-432	Sequence 432, App
450	34	64.2	346	14	US-10-152-391-432	Sequence 432, App
451	34	64.2	346	14	US-10-152-399-432	Sequence 432, App
452	34	64.2	346	14	US-10-156-848-432	Sequence 432, App
453	34	64.2	346	14	US-10-157-794-432	Sequence 432, App
					US-10-157-795-432	Sequence 432, App
					US-10-157-796-432	Sequence 432, App
					US-10-160-500-432	Sequence 432, App
					US-10-121-046-432	Sequence 432, App
					US-10-123-156-432	Sequence 432, App
					US-10-123-214-432	Sequence 432, App
					US-10-125-805-432	Sequence 432, App
					US-10-124-821-432	Sequence 432, App
					US-10-152-385-432	Sequence 432, App
					US-10-152-393-432	Sequence 432, App
					US-10-152-396-432	Sequence 432, App
					US-10-153-552-432	Sequence 432, App
					US-10-153-840-432	Sequence 432, App
					US-10-156-841-432	Sequence 432, App
					US-10-156-842-432	Sequence 432, App
					US-10-156-844-432	Sequence 432, App
					US-10-156-845-432	Sequence 432, App
					US-10-156-846-432	Sequence 432, App
					US-10-121-048-432	Sequence 432, App
					US-10-121-052-432	Sequence 432, App
					US-10-121-053-432	Sequence 432, App
					US-10-121-054-432	Sequence 432, App
					US-10-121-063-432	Sequence 432, App
					US-10-123-213-432	Sequence 432, App
					US-10-123-213-432	Sequence 432, App
					US-10-123-221-432	Sequence 432, App
					US-10-123-322-432	Sequence 432, App
					US-10-123-771-432	Sequence 432, App
					US-10-123-911-432	Sequence 432, App
					US-10-124-823-432	Sequence 432, App
					US-10-125-931-432	Sequence 432, App
					US-10-125-932-432	Sequence 432, App
					US-10-123-913-432	Sequence 432, App
					US-10-140-473-432	Sequence 432, App
					US-10-140-806-432	Sequence 432, App
					US-10-140-810-432	Sequence 432, App
					US-10-140-863-432	Sequence 432, App
					US-10-141-699-432	Sequence 432, App
					US-10-141-703-432	Sequence 432, App
					US-10-141-706-432	Sequence 432, App
					US-10-141-757-432	Sequence 432, App
					US-10-141-762-432	Sequence 432, App
					US-10-142-428-432	Sequence 432, App
					US-10-142-429-432	Sequence 432, App
					US-10-142-884-432	Sequence 432, App
					US-10-143-027-432	Sequence 432, App

## ALIGNMENTS

## RESULT 1

US-10-238-607-7  
; Sequence 7, Application US/10238607  
; Publication No. US20030118602A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: HCMV- REACTIVE T CELLS AND USES THEREFOR  
; FILE REFERENCE: 1954-398  
; CURRENT APPLICATION NUMBER: US/10/238,607  
; PRIOR FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 09/692,170  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 09/534,639  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/075,257  
; PRIOR FILING DATE: 1998-05-11  
; PRIOR APPLICATION NUMBER: US 09/021,298  
; PRIOR FILING DATE: 1998-02-10  
; PRIOR APPLICATION NUMBER: US 08/950,064  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: US 08/747,488





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2004, 13:03:22 ; Search time 34 Seconds

Title: US-10-697-055-7

Perfect score: 53

Sequence: 1 TPRVTGGGM 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications AA:

1: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/prodata/1/pubppa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/prodata/1/pubppa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/prodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/prodata/1/pubppa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep.\*

10: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep.\*

11: /cgn2\_6/prodata/1/pubppa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/prodata/1/pubppa/US10\_PUBCOMB.pep.\*

14: /cgn2\_6/prodata/1/pubppa/US10\_PUBCOMB.pep.\*

15: /cgn2\_6/prodata/1/pubppa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/prodata/1/pubppa/US10\_NEW\_PUB.pep.\*

17: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pep.\*

18: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

19: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

20: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

21: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

22: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

23: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

24: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

25: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

26: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

27: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

28: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

29: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

30: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

31: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

32: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

33: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

34: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

35: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

36: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

37: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

38: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

39: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

40: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

41: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

42: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

43: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

44: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

45: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

46: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

47: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

48: /cgn2\_6/prodata/1/pubppa/US60\_PUBCOMB.pep.\*

37	69.8	748	14	US-10-156-761-12713	Sequence 12713, A
36	67.9	54	14	US-10-156-761-13012	Sequence 13012, A
36	67.9	145	14	US-10-029-386-27829	Sequence 27829, A
36	67.9	244	14	US-10-156-761-15056	Sequence 15056, A
36	67.9	276	15	US-10-369-493-15690	Sequence 15690, A
36	67.9	276	15	US-10-369-493-16082	Sequence 16082, A
36	67.9	294	15	US-10-369-493-15324	Sequence 15324, A
36	67.9	600	15	US-10-369-493-20821	Sequence 20821, A
36	67.9	855	15	US-10-369-493-10058	Sequence 10058, A
35	66.0	330	14	US-10-156-761-9249	Sequence 9249, Ap
35	66.0	338	15	US-10-374-780A-1701	Sequence 1701, Ap
35	66.0	353	9	US-09-895-943-3	Sequence 3, Appli
35	66.0	353	9	US-09-895-593-3	Sequence 7, Appli
35	66.0	361	14	US-10-226-872-7	Sequence 2, Appli
35	66.0	370	9	US-09-895-943-2	Sequence 2, Appli
35	66.0	370	9	US-09-895-593-2	Sequence 9, Appli
35	66.0	424	9	US-09-835-684-9	Sequence 9, Appli
35	66.0	424	9	US-09-880-371-9	Sequence 14, Appli
35	66.0	424	9	US-09-879-248-14	Sequence 9, Appli
35	66.0	424	14	US-10-010-390-9	Sequence 14, Appli
35	66.0	424	15	US-10-441-736-14	Sequence 23, Appli
35	66.0	432	14	US-10-314-657-23	Sequence 7910, Ap
35	66.0	444	14	US-10-156-761-7910	Sequence 7677, Ap
35	66.0	502	14	US-10-032-585-7677	Sequence 6240, Ap
35	66.0	522	15	US-10-369-493-6240	Sequence 6241, Ap
35	66.0	522	15	US-10-369-493-6241	Sequence 2, Appli
35	66.0	524	14	US-10-227-035-2	Sequence 71, Appli
34	64.2	86	14	US-10-219-220-71	Sequence 2675, Ap
34	64.2	111	15	US-10-108-260A-2675	Sequence 262, App
34	64.2	128	14	US-10-219-220-262	Sequence 17800, A
34	64.2	148	15	US-10-369-493-17800	Sequence 173, App
34	64.2	206	14	US-10-247-671-173	Sequence 4633, Ap
34	64.2	260	9	US-09-738-626-4633	Sequence 4695, Ap
34	64.2	263	9	US-09-738-626-4695	Sequence 730, App
34	64.2	288	14	US-10-102-806-730	Sequence 197, App
34	64.2	346	9	US-09-989-722-197	Sequence 197, App
34	64.2	346	9	US-09-989-723-197	Sequence 197, App
34	64.2	346	9	US-09-989-279-197	Sequence 197, App
34	64.2	346	9	US-09-989-727-197	Sequence 197, App
34	64.2	346	9	US-09-989-731-197	Sequence 197, App
34	64.2	346	9	US-09-989-732-197	Sequence 197, App
34	64.2	346	9	US-09-991-073-197	Sequence 197, App
34	64.2	346	9	US-09-990-442-197	Sequence 197, App
34	64.2	346	9	US-09-991-163-197	Sequence 197, App
34	64.2	346	9	US-09-993-604-197	Sequence 197, App
34	64.2	346	9	US-09-990-456-197	Sequence 197, App
34	64.2	346	9	US-09-989-721-197	Sequence 197, App
34	64.2	346	9	US-09-992-598-197	Sequence 197, App
34	64.2	346	9	US-09-989-293A-197	Sequence 197, App
34	64.2	346	9	US-09-989-735-197	Sequence 197, App
34	64.2	346	9	US-09-990-444-197	Sequence 197, App
34	64.2	346	9	US-09-991-181-197	Sequence 197, App
34	64.2	346	9	US-09-989-730-197	Sequence 197, App
34	64.2	346	9	US-09-990-436-197	Sequence 197, App
34	64.2	346	9	US-09-993-687-197	Sequence 197, App
34	64.2	346	10	US-09-989-734-197	Sequence 197, App
34	64.2	346	10	US-09-997-653-197	Sequence 197, App
34	64.2	346	10	US-09-993-667-197	Sequence 197, App
34	64.2	346	10	US-09-997-428-197	Sequence 197, App
34	64.2	346	10	US-09-997-666-197	Sequence 197, App
34	64.2	346	10	US-09-990-438-197	Sequence 197, App
34	64.2	346	10	US-09-990-562-197	Sequence 197, App
34	64.2	346	10	US-09-990-711-197	Sequence 197, App
34	64.2	346	10	US-09-989-726-197	Sequence 197, App
34	64.2	346	10	US-09-998-156-197	Sequence 197, App
34	64.2	346	10	US-09-990-437-197	Sequence 197, App
34	64.2	346	10	US-09-991-157-197	Sequence 197, App
34	64.2	346	10	US-09-997-514-197	Sequence 197, App
34	64.2	346	10	US-09-997-573-197	Sequence 197, App
34	64.2	346	10	US-09-991-172-197	Sequence 197, App
34	64.2	346	10	US-09-990-726-197	Sequence 197, App
34	64.2	346	10	US-09-997-559-197	Sequence 197, App
34	64.2	346	10	US-09-997-601-197	Sequence 197, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	10	US-10-238-607-7	Sequence 7, Appli
2	53	100.0	10	US-10-405-231-7	Sequence 7, Appli
3	53	100.0	561	US-09-812-079A-2	Sequence 2, Appli
4	53	100.0	561	US-10-434-982-2	Sequence 2, Appli
5	53	100.0	579	US-10-223-538-6	Sequence 6, Appli
6	53	100.0	579	US-10-223-538-8	Sequence 8, Appli
7	48	90.6	10	US-10-238-607-8	Sequence 8, Appli
8	48	90.6	10	US-10-405-231-8	Sequence 8, Appli
9	39	73.6	285	US-10-051-874-89	Sequence 2, Appli
10	39	73.6	339	US-10-109-616-2	Sequence 8, Appli
11	39	73.6	342	US-10-051-874-86	Sequence 8, Appli
12	39	73.6	342	US-10-051-874-87	Sequence 8, Appli
13	39	73.6	665	US-10-259-194A-124	Sequence 124, App
14	38	71.7	383	US-10-374-780A-2854	Sequence 2854, Ap
15	37	69.8	313	US-10-306-762-109	Sequence 109, App